Designing a Meta Report for Hotel Customer Reviews

Author
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
g.t.mendler@student.utwente.nl

ABSTRACT

Customer reviews already represent a significant decision support for modern businesses. However, the reliability of such data is lacking. By performing a literature study and interviews with hotel managers we assemble reliability issues and possible solutions involving the use of data reporting. In this paper we present a meta report for hotel customer reviews, for the purpose of providing further insight into review credibility. We found that meta reporting can address reliability concerns and provides further data context. This contributes to improved services provided by analytical suites and benefits hotel managers specifically.

Keywords

Meta data, Customer review, Hotel, Data report, Data reliability, TripAdvisor

1. INTRODUCTION

The widespread use of the Internet has generated a large amount of data sources to gather and analyze information from. This so called Big Data has become increasingly important for businesses of any size to monetize or otherwise utilize in order to stay competitive. This era has been named Industry 4.0 [18]; It is the fourth industrial revolution, invoked through the rise of digitization.

With the inevitable mass of fake or wrong information, trust has been attributed utmost importance. This quality can be achieved by trust and reputation systems [15], including customer reviews. The importance and usefulness of such reviews has been shown extensively before [21], however there are still concerns about the reliability and representative nature of data arising from customer reviews, partly due to the trend of publishing, and distributing, fake reviews. These considerations can be addressed by reporting about metadata [24], on which basis the original data can be judged and conclusions justified [2].

In the past years, online customer reviews as a form of electronic word-of-mouth have established themselves as the most important factor influencing customers' decision making [27], having a greater impact on consumer behaviour than methods such as traditional marketing, promotional messages and information provided by product providers [9, 29, 31]. Apart from the effect on customer behaviour, online reviews additionally represent a significant decision support for modern businesses. Consequently, tourism organizations should continuously enhance their products and services based on information provided by customers [19, 28].

30th Twente Student Conference on IT Febr. 2^{nd}, 2019, Enschede, The Netherlands.

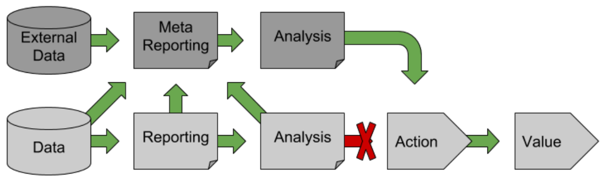

Figure 1. The data reporting workflow

Neirotti et al (2016) investigated the value created by customer reviews in the hospitality business and found that this trend is shifting hotel competition from unit profit margins to volumes and to higher room occupancy rates, which favours very large establishments with a large number of rooms. Additionally, establishments in rural areas or areas without local competition profit more from online visibility. They suggest that further research should investigate how hotels can reinforce the economic value created by the online community by enriching the gathered data [22].

The current paper aims to increase the value, credibility and usefulness of review data through meta reporting, and help hotel managers by designing the structure for a meta report. Meta reporting is a method which enriches already available information provided by traditional data reports. This can be done by offering further or more specific information about individual parts of the report, or about the report as a whole. Simple data reports can indicate the current status of an establishment, but without proper context any interpretation or analysis would not be based on the entirety of facts. Simply put, reporting uses data to track the performance of a business, whereas an analysis uses data to answer strategic questions raised by such a data report, see figure 1. With the help of a meta report, data reports can be put into perspective.

In this paper the expectations and requirements for a meta report for hotel customer review data are investigated by conducting literature research and expert interviews. On this basis an exemplary structure is constructed and presented. Some parts of this design are then visualized using a real life example, acquired through a data-scraper. This research contributes to improved services provided by analytical suites and travel agencies and shows how review data reliability can be increased.

In order to properly design a meta report, the following questions need to be addressed:

1. *How can data address reliability concerns?*
2. *How can this data be gathered and structured in a meta report?*

1.1 structure of the paper

The remainder of the paper is structured as follows. First, previous research in the fields of customer reviews and reputation systems is presented, along with an investigation of the term ”reliability”. This is followed by a description of the methodology used, which results are presented and discussed next. This section will be divided into the different sections of the meta report in no specific order, including a discussion and evaluation each. The final section contains some concluding remarks and suggestions for future research.

2. LITERATURE REVIEW

Several authors have shared their insights into metadata, data mining and customer review categorization, analysis and reliability. The concept of metadata was introduced by Prothman (2000), who also presents desirable metadata qualities like dynamicity, persistence and portability [24]. Especially Bruce et al (2004) discusses metadata quality specifically and possibilities to improve it short- and long-term [2].

The process of data mining [10] and metadata mining [23] are thoroughly discussed. For example, Jian et al (2017) performed a Big Data analysis on hotel customer reviews based on cloud computing [14] and Wan et al (2014) designed a multi-level metadata standard for museum data [26]. Girardi et al (2011) presented a general way of knowledge discovery through meta-models [8]. In August 2017 the first stable version of Selenium was released, a tool that can be used to automate data scraping for complex websites, especially those which heavily rely on JavaScript.

Additionally, several authors have done research into the field of customer reviews. Zhang et al (2010) has compared customer reviews to editorial reviews and found that customer reviews and their volume can significantly increase the popularity of a restaurant, while ratings created by editors have a negative influence [30]. Josang (2006) discusses trust and reputation systems in general, which customer reviews are a part of, and present an overview about problems and proposed solutions. These problems include, amongst other things, the bias towards a positive ratings, quality variations over time and discrimination [15].

The risk of misinformation in the internet is addressed by Viviani (2017), where he introduces the concept of credibility, described as a quality perceived by individuals not able to absolutely discern between genuine and fake information, in regard to social media and news [25]. Fake reviews can be automatically detected by using linguistic features and Latent Dirichlet Allocation, which was demonstrated successfully by Jia et al [13].

2.1 TripAdvisor as a platform

The platform TripAdvisor has itself been the focus of research. The work of Jeacle (2011) includes a case study on TripAdvisor and provides insight into the integrity and authenticity of the website [11]. Filieri (2016) assesses trustworthiness of online customer reviews and criteria for its distinction, content and writing style on the one hand, but source of communication and emerging patterns across multiple reviews on the other [6]. In another paper, Filieri discusses why travelers trust TripAdvisor as a source for ratings. This trust can be traced back to three main aspects: the quality of the information contained in online reviews, the quality of the website that hosts the recommendations, and the level of customer satisfaction with previous experiences [7].

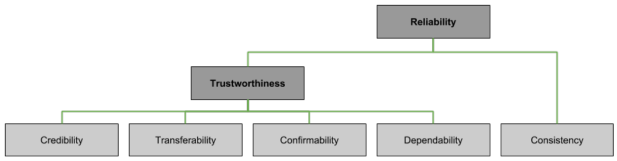

Figure 2. The composition of reliability

TripAdvisor is an online infomediation platform used by travellers to compare prices and customer reviews on different types of establishments, including hotels, restaurants and other touristic attractions. Through its widespread use, it generated value by offering market transparency and convenience for the users. Further, it mediates between online travel agencies and the establishments to enable booking through the website itself. Since TripAdvisor relies on some social media functionalities [16] to offer users to connect to other people with similar interests or experience, limited data about traveller demographic and their historical data is openly available. This can easily be connected to individual hotel reviews since a user authentication is required in order to create online reviews.

Dina Maizlin et al (2014) conducted an empirical investigation of online review manipulation, comparing TripAdvisor with Expedia, a platform on which only customers can create reviews. On TripAdvisor, hotels with a higher incentive to fake generally have more positive reviews. Moreover, hotels in the neighbourhood of competitors, which have an incentive to fake, usually have a higher amount of negative reviews [20].

2.2 Reliability as a concept

Furthermore, the concept of reliability is important, especially internal consistency reliability (Henson 2001) [3]. Generally, reliability is the combination of trustworthiness and consistency. Trustworthiness itself consists of credibility, transferability, confirmability and dependability, as can be seen in figure 2. *Consistency* describes the difference across results, scoring higher with a low difference. In order to confirm this metric it can be shown that different approaches lead to the same results, and that no two results contradict each other. *Credibility* refers to the extent to which a research account is believable and appropriate. This metric can be satisfied with statistical evaluations, for example comparing standard errors and significance intervals. *Transferability* refers to the degree to which the results of research can be generalized or transferred to other contexts or settings. This is mostly the responsibility of the one doing the generalization, however this can be supported by the researcher by stating the research context and its assumptions. Next, *confirmability*, which measures the influence of the researchers' bias on the results, this is also called degree of neutrality. Confirmability is of much higher importance in sociological studies and data sets, which rely on human interpretation to some degree, in which case it can be supported by audit trails. Lastly, *dependability* is based on the assumption of replicability.

Table 1. Scraped data fields

Category	Data field
general	title
	review text
	helpfulness
	trip type
dates	date of review
	month of stay
ratings	general rating
	room quality
	service quality
	sleep quality
	location
	cleanliness
	value
user information	profile description
	total helpfulness
	tags
	number of 5-star reviews
	number of 4-star reviews
	number of 3-star reviews
	number of 2-star reviews
	number of 1-star reviews

It describes how much data depends on special variables, therefore judging how well results can be achieved by other researchers, and the expected degree of consistency across their results.

3. METHODOLOGY

To achieve the objectives of this project, I worked according to a multi-method research design. A literature study was performed on topics such as customer reviews, TripAdvisor and data reporting. This established a knowledge base including the specific shortcomings of data reports and customer reviews, and the resulting expectations thereof.

3.1 Design process

To aid with the design process, an interview was conducted with a hotel manager. For this purpose I discussed current business practices and their flaws with a German hotel manager, who is responsible for a medium sized highly rated hotel in a rural German area called the Eifel.

The purpose of this conversation was to establish a general overview about common business practices regarding customer reviews. They were asked how they consult and judge TripAdvisor ratings and reviews, what concerns they have and what additional information they would like to know about these customer reviews.

Then, the structure of a metadata report for hotel customer reviews will be conceptualized on the basis of all previously acquired knowledge.

3.2 Data collection

For prototyping purposes an example data set was required. There are data sets available containing TripAdvisor online reviews, however all these show only essential data fields such as data, title and text. A meta report naturally required more detailed data, therefore a custom scraping tool was developed. This was done using the R language, designed for statistical computing, which first appeared in 1993. Relying on the previously mentioned Selenium library, the R script can scrape the data of all reviews for a specific given hotel. This includes general data like title, text and rating, but also user information, date of stay and subratings for service, room quality and much more. All 21 data fields are summarized in table 1. The helpfulness metric relies on an internal TripAdvisor system, which enables users to indicate that a specific review is helpful, similar to a social media like-system. This count can only be increased.

The data set acquired through this script was then further cleaned up and expanded by other software. Statistical methods were applied through Google Sheets, which was also used to generate all necessary graphs.

3.3 Evaluation

An evaluation has been carried out for every section of the meta report individually, examining the satisfaction of the previously established reliability metrics. For this purpose real life data was used. An adequate specimen needed to fulfill several requirements: The establishment should be situated in a well contested area and have a large amount of customer reviews. It should be noted that the customer base of any hotel is very limited compared to other products like cars or smartphones, therefore the number of customer reviews rarely exceeds 1,000, even in very large cities and over the course of several years.

Therefore the pod 39 hotel in the centre of New York City, a mid-tier establishment with 366 rooms, was selected. Although not the most prominent local business, it lists nearly 4,000 reviews, 2,300 of them distributed over the past 3 years. As comparisons, another hotel of equal class, the Hilton at West 35th street, and a luxury hotel, the Archer, were chosen. Both have a similar number of available customer reviews, although slightly smaller in size.

4. RESULTS

4.1 Important observations

From the interview it emerged that, in general, it can be expected that any hotel management is able to read through all customer reviews in person. Therefore the data context and overarching statistics are much more helpful than data about individual customer reviews. Additionally, substandard ratings are of higher interest to the management than beneifical reviews, which are much more common, due to a tendency to leave positive reviews. Compared to positive reviews, negative ones have a higher potential to point out flaws and criticisms. The reasons behind a low rating are especially interesting. Correlations between special circumstances and rating tendencies are essential for business decisions.

4.2 Assumptions

It is assumed that TripAdvisor is a basis for realistic customer reviews, that reflect the general opinions of travellers. This also includes the assumption that only a minor portion of reviews is fake or artificial. It is not part of this study to try to identify or remove any fake customer reviews. It should be noted that TripAdvisor already has measures in place that track user behaviour to delete fake reviews. As previously mentioned, these measures are not always effective [20, 5], however sufficiently reliable to support this hypothesis [4].

4.3 Report structure

The proposed structure for the meta report consists of multiple sections, in no specific order. All examples are for December 2018, unless otherwise stated.

Table 2. Average monthly and yearly rating

Value	Rating	Standard Error
Year 2017	4.09	0.045
Year 2018	4.28	0.046
November 2018	4.56	0.123
December 2018	4.07	0.192

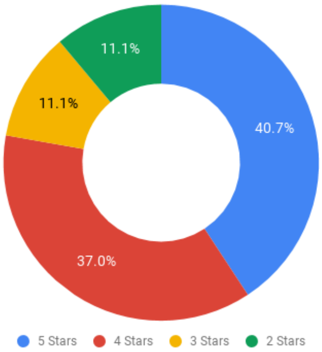

Figure 3. Rating Distribution December 2018

4.3.1 *Overview data*

The purpose of this section is to provide a general overview of the data handled in this meta report. It provides general averages and a simple context to support the formation of reasonable expectations. The general rating is the most important value for travellers, as it combines all other data into a single score on a simple one to five scale.

As can be seen in table 2, the yearly average increased significantly from 2017 to 2018 by 4.6%. However, the monthly average of December 2018 is significantly lower than the previous month, November 2018, by approximately 11.4%. Clearly, December is also considerable worse compared to the entire Year 2018, whereas November seems to have been a better than average month for this hotel.

In figures 3 and 4 it can be seen how the average ratings are constituted. During December, 1,2 and 3 star ratings form 22% of all ratings, whereas 15% is the yearly percentage. Further, the portion of 5 star ratings in December is much lower than normal.

This section has a high credibility due to significant differences between results. Also, confirmability is addressed by using statistical computations only, without any need for human interpretation. Therefore the results should be similar for all researchers attempting to repeat this study.

Table 3. Average rating per trip type

Trip type	Rating	Standard Error
with Family	4.24	0.042
Solo	4.27	0.034
with Friends	4.22	0.043
as a Couple	4.12	0.029
on Business	4.08	0.057

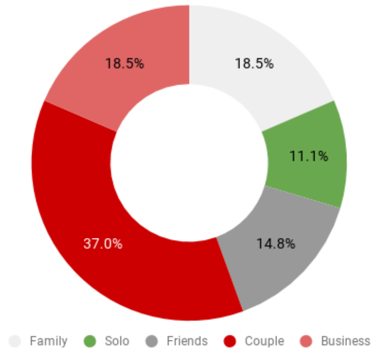

Figure 5. Trip type distribution December 2018

4.3.2 *Customer demographic*

This section investigates the customers that wrote the reviews. Different demographic information can be found on TripAdvisor. This includes age, gender and origin for some travellers, based on voluntary publication. However, gender is most likely not a reliable measure, since many users do not travel alone, therefore other members of the group would not be considered. There are not enough entries to properly support any significant hypothesis about age and origin of the customers.

With the available data the correlation between ratings and trip types can be investigated. Trip types indicate if, and which whom, the review author travelled. There are five different types, as seen in table 3. Solo travellers generally rate significantly higher than customers on business trips and couples. This is especially interesting since the Pod 39 Hotel is advertised specifically for business and romantic trips. Based on this information, the management should reconsider and focus more towards solo and family travellers.

Furthermore, trip type information can explain the below average general rating in December 2018. Comparing figure 5 and 6, it can be observed that the portion

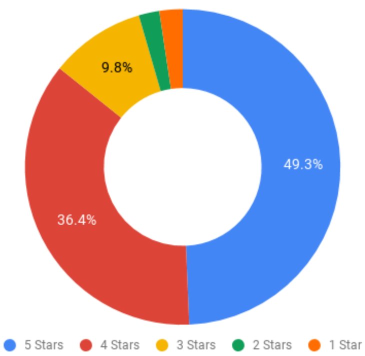

Figure 4. Rating Distribution Year 2018

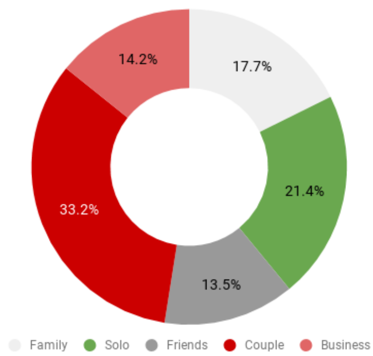

Figure 6. Trip type distribution Year 2018

of high-rating customers during December 2018 is significantly lower than the yearly average. The percentage of customers travelling alone nearly halved and the business and couple parts increased instead. In other words, there is a higher percentage of business travellers and couples than normal, who tend to leave lower ratings.

To further improve this section, additional demographic data could be scraped from social media profiles of travellers. Especially information about disabilities or personal restrictions and their experience during the stay is desired but ultimately proves hard to find.

This section addresses the consistency and credibility aspects of reliability. The results confirm the first section, and are significant enough themselves through statistical computations. Transferability is provided under the general assumption of represantativity, therefore the same results should be obtained with customer reviews sourced from other sources like Trivago or Booking.com.

4.3.3 External factors

Aside from customer demographics, other external factors can also heavily influence online customer ratings, as shown by Bakhshi et al in their 2014 study about restaurants. This could include weather data or local and national news, as well as the neighbourhood an establishment is located in. In their study they discovered that demographic and weather data can have a significant effect on customer ratings. Often, these effects can be explained using theories from experimental psychology [1].

In order to support the other sections of this meta report, this section should especially focus on correlations between external data and results acquired throughout this entire report, for example the effect of rain on monthly averages. These correlations need to be sufficiently significant to satisfy the consistency and credibility requirements.

4.3.4 Text analysis

Another potential way of acquiring interesting demographic information, but also additional new information, is text analysis. Each review on TripAdvisor includes text, which can easily be translated into English using tools like the Google translate API. In the data set acquired by the custom tool during this research, all texts are already translated into English. High-frequency words and phrases, major topics and subtopics, as well as mentioned issues, employee names and the timespan and length of visit should be identified. This can be done similarly to the approach of Susan Jia, who used multilinear regression to analyze the text of restaurant reviews in her 2018 study [12].

Results of such methods should be used to reinforce other sections, since it can be expected that all review texts have been read already by the person responsible. A helpful evaluation could focus on the correlation between text sentiment score and the given star rating, since a significant connection further proves the representative nature of the general ratings.

This section needs to take special care to satisfy the reliability requirements, especially confirmability and dependability, since sentiment analysis is usually not transparent.

4.3.5 User platform data

An entirely different approach to support previous observations is the analysis of user data gathered by TripAdvisor. Since the creation of a review required a signed in session, all customer reviews can be linked to some user account. TripAdvisor keeps track of all previous user activity for each account, for example a list of all reviews a user has written. This can be used to help identify fake reviews, but also provides a context for the actual review in question. This convenient rating history facilitates the analysis of traveller rating behaviour.

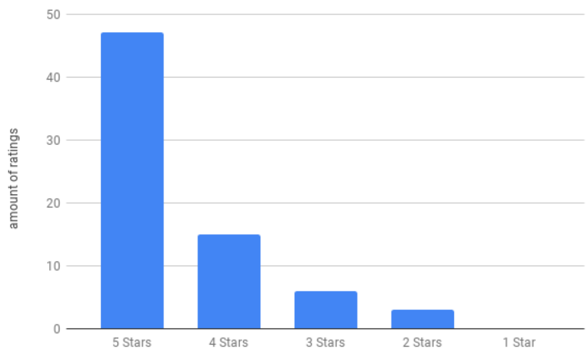

Figure 7. Historic rating distribution of example traveler

Table 4. Attitude averages

Value	Attitude	Standard Error
2015 to 2018	-0.05	0.016
Year 2017	-0.14	0.042
Year 2018	-0.02	0.043
November 2018	0.17	0.11
December 2018	-0.27	0.18

After counting the number of occurrences of each grade of ratings, the mean of these can be computed. This is shown for an exemplary user in figure 7. This value represents the average rating a traveller makes. This user has an mean rating of 4.49. When subtracting the actual rating for the establishment in question, an attitude score is acquired. This example traveller wrote a 4-Star review, therefore the computed attitude is -0.49. This score indicates the difference between the reported and perceived rating. If this score is positive, the hotel is perceived as above average for this customer, otherwise below. The reported rating can vary significantly across cultures, as shown by Noi Sian Koh and others in their 2010 investigation about movie reviews, which shows that especially Americans have a tendency towards under-reporting, meaning "consumers with extreme opinions are more likely to report their opinions than consumers with moderate reviews causing online reviews to be a biased estimator of a product's true quality". Therefore the reported rating is not a sufficient indicator for actual performance [17]. In this metric, reviews which differ greatly from their respective averages are attributed a more extreme attribute score and consequently have a higher weight and effect on general results. Moreover, users which only leave 1 star reviews or only 5 star reviews have no extreme influence on the following statistics, since their attitude score is 0.

As can be seen in table 4, the Pod 39 Hotel has been slightly below average for the past 3 years. As the first section indicated, the year 2018 has in fact been perceived as significantly better than the year 2017. In fact, the values for year 2018 even indicate a positive trend, as they are higher than the 3-year average. On a monthly basis, the previous results of section 1 are also confirmed. November 2017 has been an exceptionally good month, after which December stands out with significantly worse scores, compared to both the previous month and the entire year.

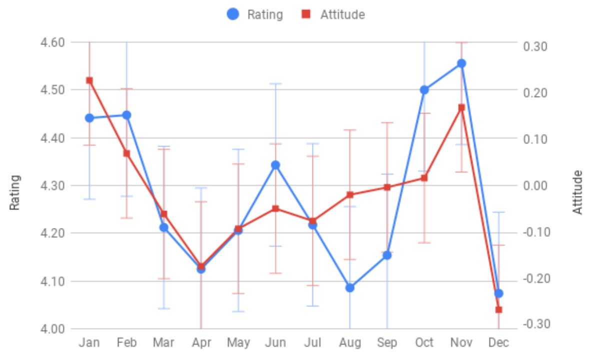

Figure 8. Comparison of monthly rating and attitude changes 2018

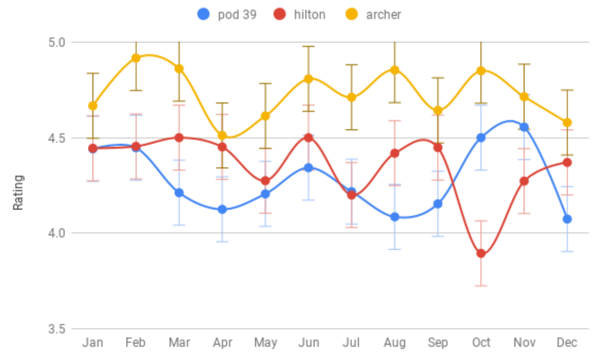

Figure 9. Comparison of monthly average ratings across hotels 2018

Table 5. Comparison with competitors 2018

Value	Pod 39	Hilton	Archer
rating	4.28	4.36	4.72
std error	0.046	0.034	0.023
user average	4.22	4.28	4.35
std error	0.010	0.018	0.018
attitude	-0.02	0.10	0.30
std error	0.043	0.049	0.034

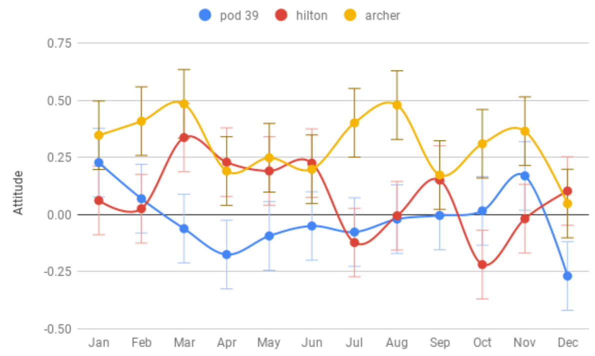

Figure 10. Comparison of monthly average attitudes across hotels 2018

In figure 8 it can be seen that the monthly attitude changes mostly reflect the trends in rating, however differ in June, August and September.

This section especially satisfies the consistency aspect, as the results confirm the data of previous sections. Credibility is also addressed, however dependability suffers from some problems. Given the same data set, repeatability is no issue, but when scraping fresh data from TripAdvisor, recent ratings of users who also review the Pod 39 Hotel will influence the attitude scores, therefore the results may change. Transferability of this sections' data is very limited, since the attitude scores will differ for each hotel, since they represent a measure of individual performance. Further, it can not be guaranteed that user averages are similar across different rating websites.

4.3.6 Comparisons

This last section is meant to provide a competitive context for the establishment in question. In this case, the Pod 39 Hotel will be compared to two other hotels in the same neighbourhood, midtown New York City. As previously mentioned, the Hilton at West 35th street is also a 3-star establishment, as is the Pod 39 Hotel, and therefore attracts similar travellers. The Archer is a luxury 4-Star Superior hotel and not a direct competitor to the other two.

The data from previous sections can be compared to these hotels to properly understand the market positioning. The overview data can be found in table 5. It emerges that the Hilton, although in the same class, is of higher quality as it attracts better reviews and has a significantly higher attitude score. Interestingly, the average user rating increases together with the average hotel rating, therefore customers of the Archer have an average rating of 3.35, significantly higher than the user average of 4.22 of the Pod 39. Comparisons on a monthly basis can be seen in figures 9 and 10.

As with the overview data it can be observed that the Archer hotel is significantly higher rated than the other two, whereas the Hilton seems to be generally higher than the Pod 39 Hotel, although this fluctuates. Additionally both the rating and attitude averages change similarly over the months, and do not contradict each other, except the values of March to June for the Hilton. Here, the attitude scores suggest similar traveller satisfaction to the Archer, whereas the rating scores are more similar to the Pod 39 Hotel.

This section fulfills the credibility requirement through significant error margins. Different results in this section do not contradict each other, therefore complying with the consistency metric.

4.4 Reliability

For each section its accordance to the previously acquired reliability metrics was discussed. The method of result acquiry is transferable itself, as it can be applied to any hotel listed on TripAdvisor. Further, all results satisfy the confirmability aspect, as no result is influenced by any researchers' bias. This is ensured through computational means, as each result relies solely on the original data set. Generally, the results have a low dependabiltiy, as they can all be achieved by other researchers, especially with the same original data set.

Naturally all statistical computations suffer from the relatively small sample size. This could be counteracted by consulting several other travel agencies and combining the data. Since customers are unlikely to rate the same trip at different websites, overlapping entries are no significant concern. This would especially increase the credibility of the results.

5. CONCLUSION

From these results it can be concluded that all 5 metrics of reliability have been successfully addressed, such that a reliable meta report was created. Further the meta report provided a context to properly place the pod 39 Hotel in, which was revealed to be below customer expectations.

This study may have practical implications for the design of analytical suites of travel websites and hotel managements and their decision making either directly or as a consequence. Further, it may lead to an improved design for online recommendation sites. Scientific implications include that all review data reliability can be improved using similar meta reports. Furthermore it was observed that the distribution of star ratings, including all historical ratings, of luxury hotel customers across the 5 star scale does not differ significantly from customers of cheaper hotels. This implies that there may be a generally valid distribution of star ratings across all travelers.

Future work might include the introduction of further data, techniques and therefore results to further reinforce the meta report. Also, the development of a user friendly application, web interface or API might benefit hotel managers or other interested parties. Additionally, the data provided by this meta report might be the basis for further research, for example into the differences of attitude scores across cultures. Moreover, the integration of proper and reliable age, origin and gender information would positively influence the demographic section especially.

6. ACKNOWLEDGMENTS

Special thanks to Mr. Dreyer for making time for our conversation. Additionally, grateful thanks to Fons Wijnhoven and his assistance and guidance during this research project.

7. REFERENCES

[1] S. Bakhshi, P. Kanuparthy, and E. Gilbert. Demographics, weather and online reviews: A study of restaurant recommendations. In *Proceedings of the 23rd international conference on World wide web*, pages 443–454. ACM, 2014.

[2] T. R. Bruce and D. I. Hillmann. The Continuum of Metadata Quality: Defining, Expressing, Exploiting. ALA Editions, 2004.

[3] M. M. Capraro, R. M. Capraro, and R. K. Henson. Measurement error of scores on the mathematics anxiety rating scale across studies. *Educational and Psychological Measurement*, 61(3):373–386, 2001.

[4] A. Y. Chua and S. Banerjee. Reliability of reviews on the internet: The case of tripadvisor. In *World Congress on Engineering & Computer Science*, pages 453–457. York, 2013.

[5] A. J. Cordato. Can tripadvisor reviews by trusted? *Travel Law Quarterly*, pages 257–263, 2014.

[6] R. Filieri. What makes an online consumer review trustworthy? *Annals of Tourism Research*, 58:46–64, May 2016.

[7] R. Filieri, S. Alguezaui, and F. McLeay. Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management*, 51:174–185, Dec. 2015.

[8] D. Girardi, J. Dirnberger, and M. Giretzlehner. Meta-model based knowledge discovery. pages 8–12, Milano, Italy, Sept. 2011. IEEE.

[9] U. Gretzel and K. H. Yoo. Use and impact of online travel reviews. *Information and communication technologies in tourism 2008*, pages 35–46, 2008.

[10] D. J. Hand, H. Mannila, and P. Smyth. *Principles of data mining*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 2001.

[11] I. Jeacle and C. Carter. In TripAdvisor we trust: Rankings, calculative regimes and abstract systems. *Accounting, Organizations and Society*, 36(4-5):293–309, May 2011.

[12] S. Jia. Behind the ratings: Text mining of restaurant customersâĂŹ online reviews. *International Journal of Market Research*, page 1470785317752048, 2018.

[13] S. Jia, X. Zhang, X. Wang, and Y. Liu. Fake reviews detection based on lda. In *2018 4th International Conference on Information Management (ICIM)*, pages 280–283. IEEE, 2018.

[14] M.-S. Jian, Y.-C. Fang, Y.-K. Wang, and C. Cheng. Big data analysis in hotel customer response and evaluation based on cloud. pages 791–795, Pyeongchang, Kwangwoon Do, South Korea, 2017. IEEE.

[15] A. JÃÿsang, R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618–644, Mar. 2007.

[16] J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre. Social media? get serious! understanding the functional building blocks of social media. *Business horizons*, 54(3):241–251, 2011.

[17] N. S. Koh, N. Hu, and E. K. Clemons. Do online reviews reflect a productâĂŹs true perceived quality? an investigation of online movie reviews across cultures. *Electronic Commerce Research and Applications*, 9(5):374–385, 2010.

[18] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann. Industry 4.0. *Business & Information Systems Engineering*, 6(4):239–242, Aug. 2014.

[19] S. E. Levy, W. Duan, and S. Boo. An analysis of one-star online reviews and responses in the washington, dc, lodging market. *Cornell Hospitality Quarterly*, 54(1):49–63, 2013.

[20] D. Mayzlin, Y. Dover, and J. Chevalier. Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8):2421–55, 2014.

[21] S. Melián-González, J. Bulchand-Gidumal, and B. González López-Valcárcel. Online customer reviews of hotels: As participation increases, better evaluation is obtained. *Cornell Hospitality Quarterly*, 54(3):274–283, 2013.

[22] P. Neirotti, E. Raguseo, and E. Paolucci. Are customersâĂŹ reviews creating value in the hospitality industry? exploring the moderating effects of market positioning. *International Journal of Information Management*, 36(6):1133–1143, 2016.

[23] A. Pomares-Quimbaya, M. E. Torres-Moreno, and F. RoldÃạn. MetaExtractor: A System for Metadata Extraction from Structured Data Sources. volume 8127, pages 84–99. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[24] B. Prothman. Meta data. *IEEE Potentials*, 19(1):20–23, Mar. 2000.

[25] M. Viviani and G. Pasi. Credibility in social media: opinions, news, and health information-a survey:

Credibility in social media. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(5):e1209, Sept. 2017.

[26] J. Wan, Y. Zhou, G. Chen, and J. Yi. Designing a Multi-level Metadata Standard based on Dublin Core for Museum Data. *International Conference on Dublin Core and Metadata Applications*, pages 31–36, Oct. 2014.

[27] Q. Ye, R. Law, B. Gu, and W. Chen. The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human behavior*, 27(2):634–639, 2011.

[28] B. Zeng and R. Gerritsen. What do we know about social media in tourism? a review. *Tourism Management Perspectives*, 10:27–36, 2014.

[29] K. Zhang, R. Narayanan, and A. N. Choudhary. Voice of the customers: Mining online customer reviews for product feature-based ranking. *WOSN*, 10:11–11, 2010.

[30] Z. Zhang, Q. Ye, R. Law, and Y. Li. The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, 29(4):694–700, Dec. 2010.

[31] F. Zhu and X. Zhang. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, 74(2):133–148, 2010.

APPENDIX

A. R CODE

A.1 Get review urls

This method gathers all or a specific amount of most recent customer reviews of a given hotel url on TripAdvisor and returns a list of their urls.

```
get_all_reviews <- function(amount = 0, url) {
   library(Rcrawler)
   library(RSelenium)

   get_review_urls <- function(url, remDr, amount) {
      remDr$navigate(url)

      name <- remDr$findElement(
         using = "xpath",
         value = "//*[@id='HEADING']"
      )$getElementText()
      rating <- remDr$findElement(
         using = "xpath",
         value =
             "//span[contains(@class,'ui_bubble_rating')]"
      )$getElementAttribute("class")

      remDr$navigate(paste(url,"#REVIEWS",sep=""))
      Sys.sleep(1)

      remDr$findElement(
         using = "xpath",
         value ="//div[@class='choices
             is-shown-at-tablet']/div[1]"
      )$clickElement();

      Sys.sleep(1)

      num <- remDr$findElement(
         using = "xpath",
         value =
             "//span[@class='reviews_header_count']"
      )$getElementText()
      num <- strtoi(gsub(",","",substring(num, 2,
          nchar(num)-1)))

      urlElems <- remDr$findElements(
         using = "xpath",
         value = "//a[span[@class='noQuotes']]")
      urlList <- unlist(
         lapply(
            urlElems,
            function(x){x$getElementAttribute("href")}))

      pages <- (num %/% 5)
      for (i in 1:pages) {
        if (amount != 0 & i > amount) {break}
        print(paste(name, i, "of", pages))
         remDr$findElement(
            using ="xpath",
            value = "//div[@class='unified
                ui_pagination ']/a[2]"
            )$sendKeysToElement(list("\uE007"))

         Sys.sleep(1)

         urlElems <- remDr$findElements(
            using = "xpath",
            value = "//a[span[@class='noQuotes']]")
         temp <- lapply(
            urlElems,
            function(x){x$getElementAttribute("href")})
         urlList <- unlist(c(temp, urlList), recursive
             = TRUE)
      }

      urlList
   }

   driver <<- rsDriver()
   remDr <<- driver[["client"]]
   remDr$setImplicitWaitTimeout(milliseconds = 10000)

   temp <- try(get_review_urls(url, remDr, amount))

   remDr$close()
   driver$server$stop()

   temp
}
```

A.2 Scrape reviews

Given a list of review urls, this method scrapes all of the 21 data fields for each element on the list, returning a list of lists as result, in the same order as the argument.

```
scrape_reviews <- function (urlList) {
   library(Rcrawler)
   library(RSelenium)

   scrape_review <- function(url, remDr, flag =
       FALSE) {
      remDr$navigate(url)
      Sys.sleep(1)
      remDr$setImplicitWaitTimeout(milliseconds = 0)
      remDr$maxWindowSize()

      if (flag) {
         remDr$setImplicitWaitTimeout(milliseconds =
             2500)
         try(remDr$findElement(
            using = "xpath",
            value = "//div[@class='rsdc-title']"),
            silent = TRUE)
         remDr$findElement(
            using = "xpath",
            value = "//h1[contains(@class,'header')]"
            )$clickElement()
         remDr$setImplicitWaitTimeout(milliseconds = 0)
      }

      more_button <- try(remDr$findElement(
         using = "xpath",
         value = "//span[contains(@class,'moreBtn')]"))
      if(class(more_button) != "try-error") {

         try(more_button$clickElement())
         Sys.sleep(1)
      }

      #text / title
      trans_button <- try(remDr$findElement(
         using = "xpath",
         value =
             "//div[@class='featured-review-container']//span[text()
             = 'Google Translation ']"))
      if (class(trans_button) != "try-error") {
         remDr$setImplicitWaitTimeout(milliseconds =
             10000)
         trans_button$clickElement()
         title <- remDr$findElement(
            using = "xpath",
            value = "//div[@class='review userreview
                translatedoverlay']/div[1]"
            )$getElementText()
         text <- remDr$findElement(
            using = "xpath",
            value = "//div[@class='review userreview
                translatedoverlay']/div[3]"
            )$getElementText()
         remDr$setImplicitWaitTimeout(milliseconds = 0)
         remDr$sendKeysToActiveElement(list(key =
             "escape"))
      } else {

         title <- remDr$findElement(
            using = "xpath",
            value = "//h1[@id='HEADING']"
```

```
      )$getElementText()
   text <- remDr$findElement(
      using = "xpath",
      value =
           "//span[contains(@class,'fullText')]"
      )$getElementText()
}

remDr$setImplicitWaitTimeout(milliseconds = 0)

#stay date
stay_date <- remDr$findElement(
   using = "xpath",
   value =
        "//div[span[@class='stay_date_label']]"
   )$getElementText()
stay_date <- substring(stay_date,15)
#trip type
trip_type <- remDr$findElement(
   using = "xpath",
   value =
        "//div[span[@class='trip_type_label']]"
   )$getElementText()
trip_type <- substring(trip_type,12)
#value rating
val_rating <- try(remDr$findElement(
   using = "xpath",
   value = "//li/div[contains(@class,
        'ui_bubble_rating') and
        ../div[@class='recommend-description'
        and contains(text(),'Value')]]"
   )$getElementAttribute("class"),
   silent = TRUE)
val_rating <-
     as.numeric(substring(val_rating,25))/10
#room rating
room_rating <- try(remDr$findElement(
   using = "xpath",
   value = "//li/div[contains(@class,
        'ui_bubble_rating') and
        ../div[@class='recommend-description'
        and contains(text(),'Rooms')]]"
   )$getElementAttribute("class"),
   silent = TRUE)
room_rating <-
     as.numeric(substring(room_rating,25))/10
#location rating
loc_rating <- try(remDr$findElement(
   using = "xpath",
   value = "//li/div[contains(@class,
        'ui_bubble_rating') and
        ../div[@class='recommend-description'
        and contains(text(),'Location')]]"
   )$getElementAttribute("class"),
   silent = TRUE)
loc_rating <-
     as.numeric(substring(loc_rating,25))/10
#cleanliness rating
clean_rating <- try(remDr$findElement(
   using = "xpath",
   value = "//li/div[contains(@class,
        'ui_bubble_rating') and
        ../div[@class='recommend-description'
        and contains(text(),'Cleanliness')]]"
   )$getElementAttribute("class"),
   silent = TRUE)
clean_rating <-
     as.numeric(substring(clean_rating,25))/10
#sleep quality rating
slq_rating <- try(remDr$findElement(
   using = "xpath",
   value = "//li/div[contains(@class,
        'ui_bubble_rating') and
        ../div[@class='recommend-description'
        and contains(text(),'Sleep Quality')]]"
   )$getElementAttribute("class"),
   silent = TRUE)
slq_rating <-
     as.numeric(substring(slq_rating,25))/10
#service rating
```

```
serv_rating <- try(remDr$findElement(
   using = "xpath",
   value = "//li/div[contains(@class,
        'ui_bubble_rating') and
        ../div[@class='recommend-description'
        and contains(text(),'Service')]]"
   )$getElementAttribute("class"),
   silent = TRUE)
serv_rating <-
     as.numeric(substring(serv_rating,25))/10
#rating date
rat_date <- remDr$findElement(
   using = "xpath",
   value = "//span[@class='ratingDate']"
   )$getElementAttribute("title")
#helpfulness
helpf <- remDr$findElement(
   using = "xpath",
   value = "//span[@class='numHelp
        emphasizeWithColor']"
   )$getElementText()
if(nchar(helpf)>0) {helpf <- as.integer(helpf)}
#general rating
gen_rating <- try(remDr$findElement(
   using = "xpath",
   value = "//div/span[contains(@class,
        'ui_bubble_rating') and
        ../span[@class='ratingDate']]"
   )$getElementAttribute("class"))
gen_rating <-
     as.numeric(substring(gen_rating,25))/10

remDr$setImplicitWaitTimeout(milliseconds = 2500)

#USERINFO
hov_elem <- try(remDr$findElement(
   using = "xpath",
   value = "//div[@class='memberOverlayLink']"))
try(remDr$mouseMoveToLocation(webElement=hov_elem))
#user description
user_descr <- try(remDr$findElement(
   using = "xpath",
   value =
        "//ul[@class='memberdescriptionReviewEnhancements']/li[
   )$getElementText())

remDr$setImplicitWaitTimeout(milliseconds = 0)

#helpful votes
user_helpf <- try(remDr$findElement(
   using = "xpath",
   value = "//span[contains(text(),'Helpful
        votes')]"
   )$getElementText(),
   silent = TRUE)
user_helpf <-
     as.integer(substring(user_helpf,1,nchar(user_helpf)-14))
#user excellent reviews
user_exce <- try(remDr$findElement(
   using = "xpath",
   value = "//div[contains(@class,
        'histogramReviewEnhancements')]/ul/div[1]/span[3]"
   )$getElementText(),
   silent = TRUE)
user_exce <- as.integer(user_exce)
#user very good reviews
user_vego <- try(remDr$findElement(
   using = "xpath",
   value = "//div[contains(@class,
        'histogramReviewEnhancements')]/ul/div[2]/span[3]"
   )$getElementText(),
   silent = TRUE)
user_vego <- as.integer(user_vego)
#user average reviews
user_avg <- try(remDr$findElement(
   using = "xpath",
   value = "//div[contains(@class,
        'histogramReviewEnhancements')]/ul/div[3]/span[3]"
   )$getElementText(),
   silent = TRUE)
```

```
    user_avg <- as.integer(user_avg)
    #user poor reviews
    user_poor <- try(remDr$findElement(
      using = "xpath",
      value = "//div[contains(@class,
           'histogramReviewEnhancements')]/ul/div[4]/span[3]"
      )$getElementText(),
      silent = TRUE)
    user_poor <- as.integer(user_poor)
    #user terrible reviews
    user_terr <- try(remDr$findElement(
      using = "xpath",
      ^value = "//div[contains(@class,
           'histogramReviewEnhancements')]/ul/div[5]/span[3]"
      )$getElementText(),
      silent = TRUE)
    user_terr <- as.integer(user_terr)
    #user tags
    user_tags <-
        try(unlist(lapply(remDr$findElements(
      using = "xpath",
      value =
           "//ul[@class='memberTagsReviewEnhancements']/li/a"),
      function(x){x$getElementText()})),
      silent = TRUE)

    if(!is.null(user_tags)) {
      user_tags <- paste(user_tags, collapse=", ")
    } else {
      user_tags <- NA
    }

    res <-
        list(unlist(title),gen_rating,unlist(text),
      stay_date,helpf,unlist(rat_date),trip_type,val_rating,
      serv_rating,clean_rating,slq_rating,room_rating,loc_rating,
      unlist(user_descr),user_helpf,user_exce,user_vego,user_avg,
      user_poor,user_terr,user_tags)

    for (e in res) {
      if (class(e) == "try-error") {e <- NA}
    }

    res
}

scrape_all_reviews <- function(urlList, remDr) {
    lst <- list()
    for (i in 1:length(urlList)) {
      print(paste("iteration",i,"of",length(urlList)))
      res <- scrape_review(urlList[i], remDr, i==1)
      lst[[i]] <- res
    }
    lst
}

driver <<- rsDriver()
remDr <<- driver[["client"]]
    temp <- try(scrape_all_reviews(urlList, remDr))

    remDr$close()
    driver$server$stop()

    temp
}
```

A.3 Data preperation

These methods are used to convert the previous result to a data frame from the data.frame library. This is the object that Rapidminer can work with internally.

```
prepare_data <- function(resultList) {
    df <- data.frame(t(sapply(resultList,c)))

    for (i in 1:length(resultList)) {
      if (class(df[[5]][[i]]) == "list") {df[[5]][[i]]
          <- 0}
      if (class(df[[14]][[i]]) == "try-error")
          {df[[14]][[i]] <- NA}
      if (is.na(df[[8]][[i]])) {df[[8]][[i]] <- 0}
      if (is.na(df[[9]][[i]])) {df[[9]][[i]] <- 0}
      if (is.na(df[[10]][[i]])) {df[[10]][[i]] <- 0}
      if (is.na(df[[11]][[i]])) {df[[11]][[i]] <- 0}
      if (is.na(df[[12]][[i]])) {df[[12]][[i]] <- 0}
      if (is.na(df[[13]][[i]])) {df[[13]][[i]] <- 0}

      if (is.na(df[[15]][[i]])) {df[[15]][[i]] <- 0}
      if (is.na(df[[16]][[i]])) {df[[16]][[i]] <- 0}
      if (is.na(df[[17]][[i]])) {df[[17]][[i]] <- 0}
      if (is.na(df[[18]][[i]])) {df[[18]][[i]] <- 0}
      if (is.na(df[[19]][[i]])) {df[[19]][[i]] <- 0}
      if (is.na(df[[20]][[i]])) {df[[20]][[i]] <- 0}
    }

    names(df)=c("title","rating","text","staydate","helpfulnes","rati
      "triptype","value","service","cleanliness","sleepquality","roo
      "location","userdescription","userhelpfulnes","excellent","goo
      "average","poor","terrible","usertags")

    df
}

print_woText <- function(df, filename) {
    temp <- df
    temp$text <- NULL
    temp <- as.matrix(temp)
    write.table(
      temp,
      paste(filename,"txt",sep="."),
      sep="\t",
      col.names = FALSE,
      row.names = FALSE)
}
```