

## Reviewer #1 Response

Thank you for your review observations. I will quote each observation and respond to it while indicating my corrections in the text. Aside from these corrections, we also performed a general grammar check and proofread to the new manuscript, including unchanged sections. Because of comments by reviewers, we redid the experiment with more balanced data, so the results shown are different in their quantitative values. However, our conclusions and hypothesis testing remained largely unchanged by this. We also added a new research rationale and focus, which meant new hypotheses were added to the manuscript. The new manuscript focuses on how these populations interact with the high-standard spirit of Japanese hospitality and service called Omotenashi. In this environment, will the cultural differences be strong enough that their satisfaction stems from different factors?, or will both populations be largely satisfied by service and staff related terms?

### 1. Machine learning

While this manuscript contains machine learning in its title, there is no major discussion about it. What is machine learning? How do the authors define it? Is data mining part of machine learning? If so, how so? Why is important or useful to use machine learning particularly for this study? What do the authors want to achieve with machine learning?

Machine learning and data mining often overlap, which is why we had an oversight in defining both of them and explaining their differences. In a journal about Information Technology, we assumed it was general knowledge, but we added it for consistency and for a more widespread reader understanding. In the section previously titled 3.4 "Data mining, knowledge discovery and sentiment analysis", thus renamed 3.4 "Data mining, machine learning, knowledge discovery and sentiment analysis" we introduced two paragraphs defining machine learning, its importance of use for analyzing larger samples, and what we can achieve with it. However the manuscript title changed from another reviewers request.

Also, the manuscript can be improved by justifying and elaborating on the necessity of a large sample. Why is it so important about analysing a large sample? What problem did a "small" sample used in existing literature cause? I recommend the authors to prepare a section dedicating to these questions.

In regards to the sample size, a paragraph in previously 3.2, now section 3.3 "Chinese and Western tourist behavior" already describes the importance of these bigger samples. We added a bit of text to expand on this importance.

"Although the studies focusing only on Chinese tourists or only on Western tourists have a narrow view, their theoretical contributions are valuable. We

can see that depending on the study and the design of questionnaires, as well as the destinations, the results can vary greatly. Not only that, but while there seems to be some overlap in most studies, some factors are completely ignored in one study but not in the other. However, since our study uses data mining, the definition of each subject or factor is left for the hotel customers to decide en masse via their reviews (instead of being defined by the questionnaire). This means that the factors will be selected through statistical methods alone. This opened opportunities to find factors that the writers of this study would not have contemplated, or avoid enforcing a factor on the mind of reviewers by presenting them with a question that they didn't think of by themselves. This large variety of opinions in a big sample, added to the automatic findings of statistical text analysis methods, is what gives an advantage to our study in comparison to others with smaller samples. In addition, this study could help us analyze the satisfaction and dissatisfaction factors cross-culturally and compare them with the existing literature."

2. Theoretical discussion on culture I also recommend the authors to develop the story on culture with theoretical perspectives. Why would Westerners show different preferences from Chinese? Why do Westerners value soft attributes (or managerial factors) like staff more while Chinese value hard attributes (or environmental factors) like size? Why do both Westerners and Chinese share dissatisfactory attributes while they show different satisfactory attributes (see Table 6)? While this study did not measure "culture" per se, I believe the authors can find theoretical explanations.

Another reviewer pointed out this, as well as suggesting a new research rationale. Because of it, we added new content to the introduction, changed the abstract and titles, as well as adding a new section in the discussion with theoretical perspectives on the context of Japanese hospitality. It is right that our study did not measure "culture", so we can't be sure of the explanations to certain aspects, although we do explain in the literature review that it has been previously studied how different backgrounds and cultures result in different expectations and as such, different satisfaction factors.

In summary, our expanded discussion is about the differences between Chinese and Western tourists in the context of Japanese hospitality. Western tourists had hospitality and service related satisfaction keywords at the top of their priorities. However, Chinese tourists did not. We theorize this can be from two different causes. One theory is that the Chinese customers have different expectations based on their culture background, and that these expectations lead to different reactions. Specifically, not expecting something and then getting it could be less satisfying than getting an expectation met. Another theory is that despite being in the environment of Omotenashi, their language barriers are enough to disregard hospitality because of their unease or disappointment with the staff and communication availability.

We also expanded on the discussion section by adding more to the previously 7.5, now section 7.2 Managerial vs. Environmental satisfaction, as well as add reasoning to each discussion section as well. In this section we added potential ramifications of the distribution of these factors. While Westerners can be satisfied even more by improving on managerial factors, a big portion of Chinese satisfaction comes from environmental factors. This means that it could be harder for managers to achieve an improvement in their satisfaction. However, we also claim that improving on cleanliness and making hotels tobacco-free could be a solution that benefits both populations. If this is done in addition to improving Chinese language services, both populations can be satisfied further.

We added a section 7.3 Implications for hotel managers, where we discuss different strategies that managers could follow to improve on customer satisfaction for these populations. We propose two possible strategies, which are appropriate or not depending on a hotel business model. A hotel can aim to increase the numbers of Chinese customers by investing on Chinese language services, which are a larger population but spend less percentage of their budget on hotels. On the other hand, a hotel can aim to increase cleanliness and comfort, which are bigger investments, but would satisfy Western customers, who spend almost half their budget on hotels.

As for the comment "Why do both Westerners and Chinese share dissatisfactory attributes while they show different satisfactory attributes (see Table 6)?" : They are both different. RBO results and our hypothesis testing showed that both satisfactory and dissatisfactory attributes are different for both cultures. Still, I understand that explaining the reasoning behind these results could improve the discussion.

3. Restructuring I found the discussion section (7. Discussion) misrepresents. The section is mainly reporting the findings without in-depth discussion (e.g., discussion on the reason behind such findings). I recommend combining Sections 5 and 6 into one section as an analysis section; and rename 7 Discussion as 6 Results. As suggested in the earlier section, the authors can add a new discussion section with a theoretical discussion. The current 6.1 section can be revised by presenting one-to-one or item-by-item comparison between Chinese and English-speaking customers. The current format is difficult to follow to see which attributes are similar or different between these two groups.

We added most of what was once in the discussion section to the results section, and added more in-depth discussion in section 7 as explained above. The subsections that moved up to the Results section were 6.2 Chinese tourists - A big and clean space, 6.3 Western tourists - A friendly face, and absolutely clean, 6.4 Tobacco, what's that smell?, and 6.5 Location, location, location. The price section was deleted since it was short enough to be discussed elsewhere.

We revised the tables in what previously was 6.1 to present them side to side for better following the differences in keyword ranking. They are now tables 5 and 6.