A study of the determinants of the count of customer reviews on the Airbnb platform using clustering and count regression techniques

Abstract

The *sharing economy* also known as *collaborative consumption* has massively grown in the last few years and has posed a serious threat to traditional business models. Airbnb is an example of a global company that makes sharing economy possible and gives the traditional hotels a stiff competition. However, not many scholarly articles have discussed the factors that lead to successful bookings of Airbnb properties. To address this gap in existing literature, we were encouraged to put forward an analytical framework to identify such explicit and implicit factors. In this study, we have applied Poisson, Quasi-Poisson, and Negative Binomial regressions on a dataset that consists of Airbnb properties from ten different cities across the world, where the count of customer reviews posted by guests was taken as a proxy for successful bookings. We first carried out a cluster analysis based on the number of properties to generate homogeneous "cluster cities" and regression was performed thereafter for each cluster. Among the host-generated features, *superhost*, *bedrooms*, *amenities*, and *host duration*, turned out to be statistically significant. Among user- generated features, *negative sentiments*, and *overall review scores* were found to be statistically significant. We also found that the impact of host-generated content on the count of customer reviews was significantly moderated by "superhost" badge. Consequently, guests paid a higher "price per night" for "superhost" properties, while overlooking important attributes such as "website features." Through these "cluster-specific" recommendations, our study reveals how powerful is the application of computational statistics in business decision making. Finally, sensitivity analysis was performed to check for the robustness of the determinants.

Keywords: Airbnb, Poisson, Quasi-Poisson, Negative Binomial, Text analytics