Research papers in information systems management Band 17

Daniel Schlesinger | Jonas Repschläger | Rüdiger Zarnekow

Analyse der Wohnungsbewertungen von Airbnb



Universitätsverlag der TU Berlin

Technische Universität Berlin

Daniel Schlesinger | Jonas Repschläger | Rüdiger Zarnekow

Analyse der Wohnungsbewertungen von Airbnb

Die Schriftenreihe *Research papers in information systems management* wird herausgegeben von:
Prof. Dr. Rüdiger Zarnekow

Research papers in information systems management | 17

Analyse der Wohnungsbewertungen von Airbnb

Daniel Schlesinger
Jonas Repschläger
Rüdiger Zarnekow

Universitätsverlag der TU Berlin

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.dnb.de abrufbar.

Universitätsverlag der TU Berlin, 2016
http://verlag.tu-berlin.de

Fasanenstr. 88, 10623 Berlin
Tel.: +49 (0)30 314 76131 / Fax: -76133
E-Mail: publikationen@ub.tu-berlin.de

Diese Veröffentlichung – ausgenommen Umschlagfoto –
ist unter der CC-Lizenz CC BY lizenziert.
Lizenzvertrag: Creative Commons Namensnennung 4.0
http://creativecommons.org/licenses/by/4.0/

Umschlagfoto:
Till Gregor Schröder | Florence In Black & White | 2014

Satz/Layout: Daniel Schlesinger

ISBN 978-3-7983-2844-0 (online)

ISSN 2191-639X (online)

Zugleich online veröffentlicht auf dem institutionellen Repositorium
der Technischen Universität Berlin:
DOI 10.14279/depositonce-5404
http://dx.doi.org/10.14279/depositonce-5404

Inhaltsverzeichnis

Abstract 2
Einleitung 3
Studienumfang und -methodik 4
Studienergebnisse 4
Häufigkeitsverteilungen 4
Korrelationen der Bewertungskriterien 8
Diskussion 9
Zusammenfassung 12
Quellverzeichnis 13

Abstract

Analyse der Wohnungsbewertungen von Airbnb
Wie lassen sich schlecht bewertete Wohnungen auf Airbnb identifizieren?

Mit Millionen von Nutzern bietet Airbnb im Rahmen der Share-Economy die Möglichkeit Unterkünfte von Privat zu Privat zu vermieten. Entscheidend für das Zusammenfinden von Mieter und Vermieter ist das gegenseitige Vertrauen, welches auf Airbnb durch gegenseitige Bewertungen der Nutzer und ihrer Wohnungen, sowie Verifikationsnachweisen zustande kommen soll.

Um die Frage zu klären, ob es auf Airbnb überhaupt schlecht bewertete Wohnungen gibt und welchen Einfluss die Bewertungen auf das Nutzerverhalten haben, führen wir eine statistische Auswertung mittels der Daten von einer Million Airbnb-Nutzern durch. Wir evaluieren die Häufigkeitsverteilungen einzelner Bewertungsfaktoren und überprüfen diese auf mögliche Korrelationen.

Mit Hilfe der Ergebnisse wird im Anschluss erklärt, warum vermeintlich gute Bewertungen unterdurchschnittlich sein können, ein hohes Maß an Sauberkeit für einen erfolgreichen Vermieter wichtig ist und auf welche Faktoren Gäste bei der Auswahl ihrer Unterkunft besonders achten sollten, um gute von schlechten Wohnungen zu unterscheiden.

Analysis of room-ratings on Airbnb
How to identify poor rated rooms on Airbnb?

With millions of registered users Airbnb conciliates rooms from C2C in the share economy. To get landlords and tenants together it is crucial that they trust each other. Therefore, users can verify their personality, write comments about each other and the tenants can rate the rooms they rented.

With a dataset of more than one million profiles we evaluate statistically how to distinguish between ratings in detail. Therefore, we analyze frequency distributions of single judging criteria and further identify possible correlations between them.

Finally, we present key factors for the most successful landlords on Airbnb, resume how trust indicators influence the demand of the rooms and if a good room can be identified by the ratings it received only.

Einleitung

Das 2008 gegründete Unternehmen Airbnb stellt mit – nach eigenen Angaben über zwei Millionen Inseraten weltweit – die größte Share-Economy Plattform im Bereich der kommerziellen Vermietung von Unterkünften dar (Airbnb Ireland, private unlimited company 2016). Dabei scheinen die Wertungen der Wohnungen durchweg positiv zu sein, wodurch Nutzer, die auf der Suche nach einer Unterkunft sind, mitunter keine Möglichkeit sehen, empfehlenswerte Wohnungen nach den genutzten Sterne-Wertungen effektiv zu filtern. Ziel dieser Studie ist es, die Bewertungsfaktoren für Wohnungen statistisch auszuwerten. Mit Hilfe von Mittelwerten, Standardabweichungen und Korrelationen werden gute von schlecht bewerteten Wohnungen, in dem schmalen Bereich ausschließlich positiver Bewertungen, unterscheidbar gemacht. Gäste können die Ergebnisse als Hilfsmittel für ihre künftige Suche verwenden, wohingegen Gastgeber sich im Vergleich besser selbst einschätzen können.

Studienumfang und -methodik

Für diese Studie wurden die Daten der ersten eine Millionen Nutzern analysiert, die sich zwischen März 2008 und August 2011 auf der Share-Economy Plattform Airbnb registriert haben und zur Zeit der Auswertung (November 2015) 17 609 Wohnungen bereitgestellt haben. Die Nutzer sind weltweit registriert, ein regionaler Fokus in der Auswertung erfolgt nicht. Im Fokus der Untersuchung stehen die Bewertungen der Wohnungen mittels Sternen, die statistisch mittels SPSS ausgewertet wurden. Zuvor wurden dafür die Datenbänke von doppelten Einträgen bereinigt.

Studienergebnisse

Häufigkeitsverteilungen

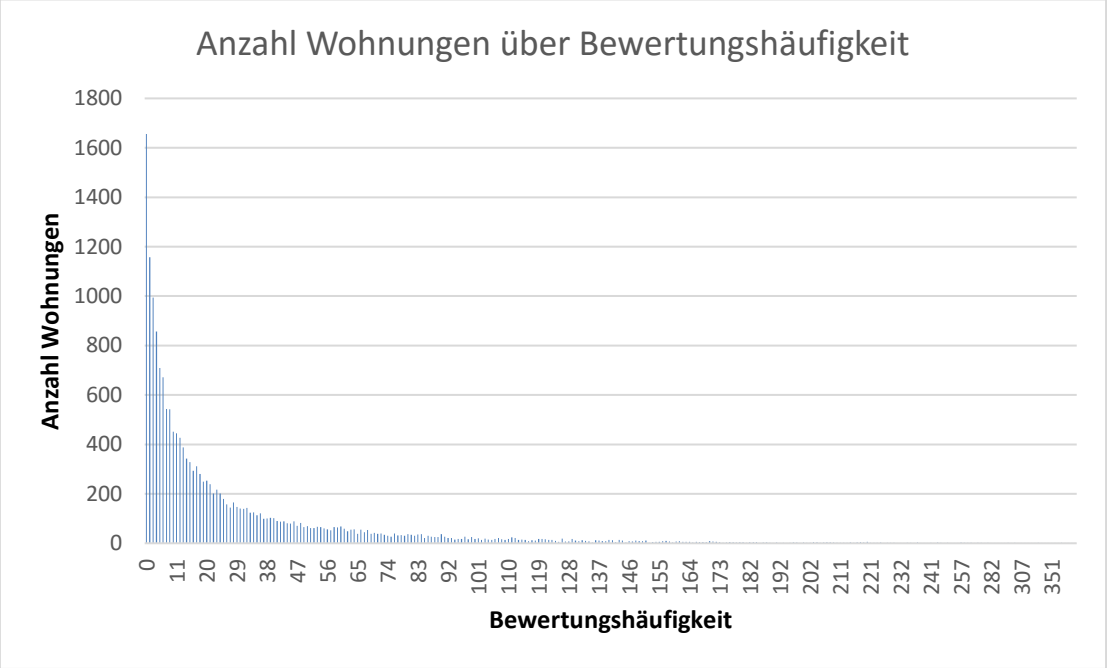


Abbildung 1: Häufigkeitsverteilung Bewertungen

Von den 17 609 Wohnungen sind 1 657 (9,4 %) weniger als drei Mal bewertet worden und nicht in die folgenden Statistiken eingeflossen, da auf Airbnb zunächst drei Bewertungen notwendig sind, bevor diese veröffentlicht werden (Airbnb Ireland, private unlimited company 2016b). Die übrigen Wohnungen wurden 492 333 mal gebucht bzw. bewertet. Nach einem Aufenthalt haben Gast und Gastgeber 14 Tage lang die Möglichkeit, sich gegenseitig mittels Freitextkommentaren zu bewerten (Airbnb Ireland, private unlimited company 2016a). Eine Wohnung wurde im Durchschnitt 27,9 Mal bewertet (vgl. Abbildung 1). Darüber hinaus hat der Gast die Möglichkeit die gebuchte Wohnung nach sechs verschiedenen Faktoren auf einer Skala von 1,0–5,0 Sternen in 0,5 Abständen einzustufen, wobei 5,0 Sterne das Optimum darstellt. Zusätzlich zu den sechs Bewertungskriterien kann der Gast eine Gesamtbewertung abgeben, welche zur Auswertung zwar nicht zur Verfügung steht, von (Zervas et al. 2015) aber ausführlich untersucht wurde. Die Bewertungskriterien sind im Folgenden:

Richtigkeit der Angaben: Hiermit urteilt der Besucher darüber, in wie weit die Unterkunft seinen Erwartungen entsprochen hat. Hierzu gehören die Fotos von

der Wohnung, die Beschreibungen seitens des Gastgebers bzgl. der Ausstattung, Hausregeln etc. (Airbnb Ireland, private unlimited company 2016c)

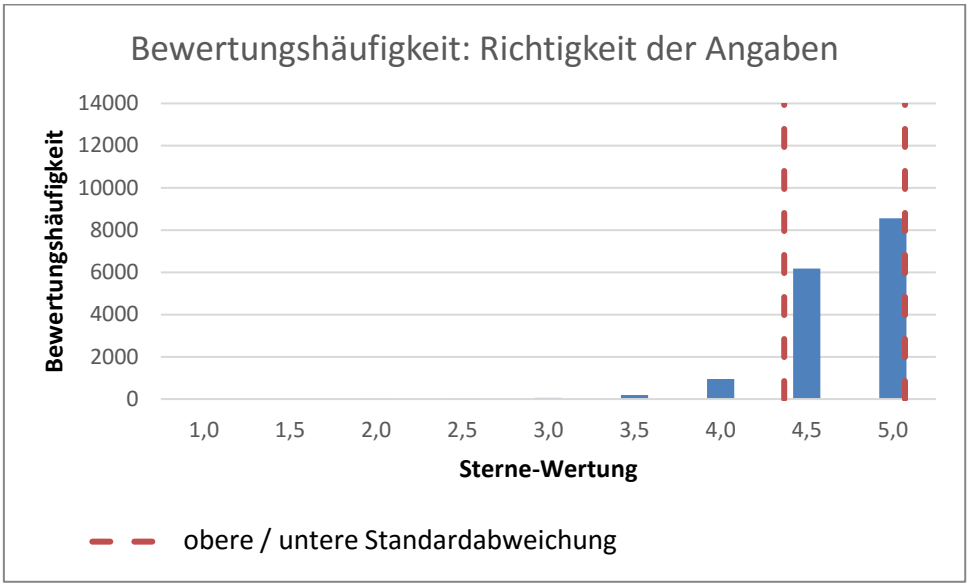


Abbildung 2: Häufigkeitsverteilung Richtigkeit der Angaben

Abbildung 2 zeigt die Häufigkeitsverteilung zur Richtigkeit der Angaben. Der Mittelwert liegt bei 4,72, Varianz und Standardabweichung bei 0,12 respektive 0,35. Damit liegen 92,4 % aller Sterne-Bewertungen von 4,5–5,0 innerhalb der Standardabweichung.

Kommunikation: Damit es zu einer Buchung kommt, müssen Besucher und Gastgeber sich verständigen, dies beginnt bei der Buchungsanfrage über den Check-In bis zur Abgabe des Wohnungsschlüssels. Die Kommunikation umfasst sämtliche Interaktionsschritte, die z. B. per SMS oder Mail durchgeführt werden. Auch der persönliche Kontakt zwischen Gastgeber und Gast wird hierbei berücksichtigt. Des Weiteren ist über die Gastgeber zu sehen, ob sie auf Anfragen reagieren (Antwortrate) und wie lange sie für eine Antwort im Schnitt benötigen (Antwortzeit). Letztere sind zwar nicht direkter Bestandteil der Bewertung, können vom Gast aber als Maßstab herangezogen werden. (Airbnb Ireland, private unlimited company 2016b)

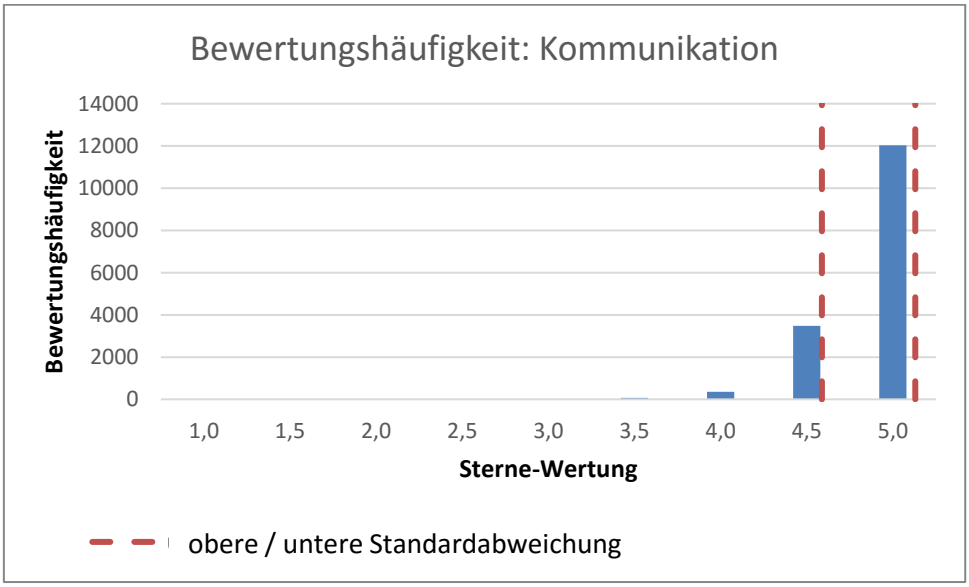


Abbildung 3: Häufigkeitsverteilung Kommunikation

Im Schnitt wird die Kommunikation mit 4,86 bewertet. Die Standardabweichung ist mit 0,27 unter allen Bewertungsfaktoren die geringste. Wie in Abbildung 3 zu erkennen liegen damit auch nur die 5,0 Sterne-Bewertungen (75 %) innerhalb der Standardabweichung.

Sauberkeit: Die Erwartungshaltung an die Sauberkeit wird auf Airbnb durch die Fotos der Wohnung, der Freitextbewertungen vorheriger Gäste, sowie der vergebenen Sterne geprägt. Abbildung 4 zeigt die Häufigkeitsverteilung der Sauberkeit. Der Mittelwert liegt bei 4,66, die Standardabweichung ist mit 0,42 unter allen Bewertungskriterien am größten. Dennoch liegen wie bei der Korrektheit der Angaben lediglich die Sterne-Bewertungen von 4,5 und 5,0 (86,2 %) innerhalb der Standardabweichung. (Airbnb Ireland, private unlimited company 2016b)

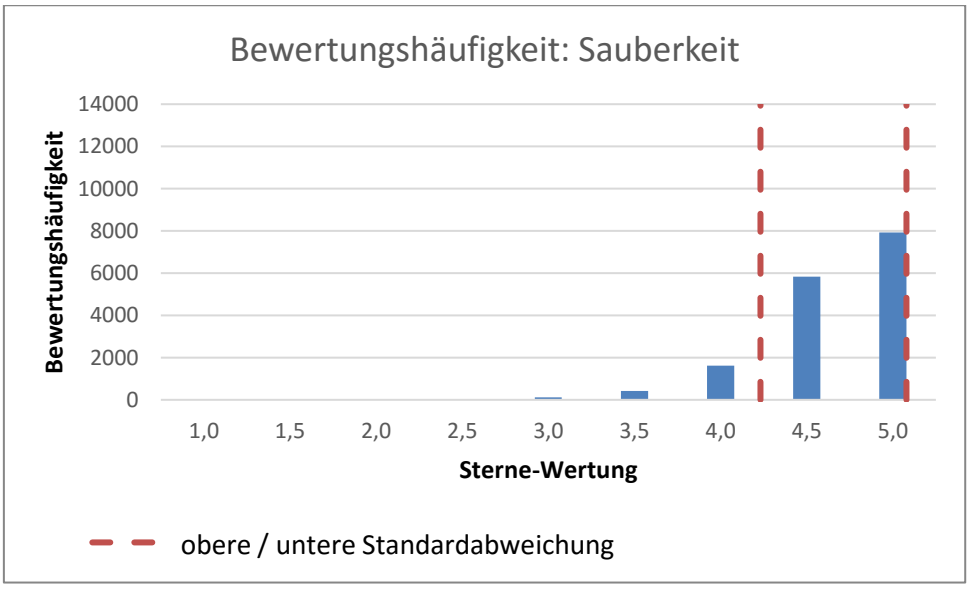


Abbildung 4: Häufigkeitsverteilung Sauberkeit

Lage: Neben Bewertungskriterien, die nur die Wohnung selbst betreffen, wird auf Airbnb auch die Umgebung bewertet. Hiermit soll sich der Gast frühzeitig eine Vorstellung von der Umgebung der Wohnung machen können. (Airbnb Ireland, private unlimited company 2016b)

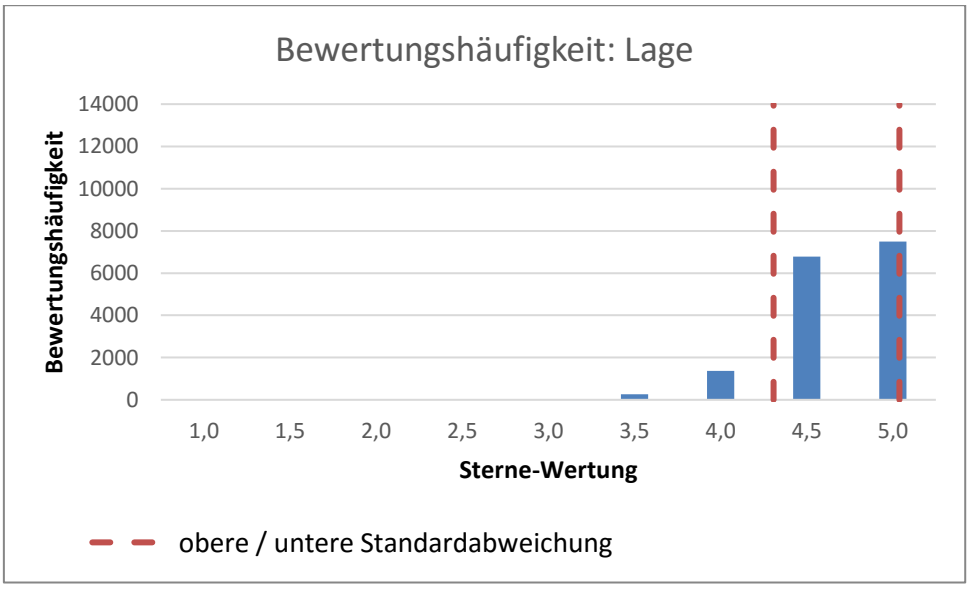


Abbildung 5: Häufigkeitsverteilung Lage

Nach der Sauberkeit ist die Standardabweichung der Lage mit 0,37 am zweithöchsten. Der Mittelwert liegt bei 4,67, so dass auch hier alle Sterne-Bewertungen von 4,5–5,0 (89,5 %) innerhalb der Standardabweichung liegen (vgl. Abbildung 5).

Check-In: Der Check-In ist Teil der Gesamtkommunikation in einem relativ zum Aufenthalt kurzem Zeitfenster. Laut Airbnb wird hier bewertet, inwieweit die Schlüsselübergabe funktioniert hat, offene Fragen geklärt werden konnten, der Gastgeber zur richtigen Zeit am richtigen Ort erschienen ist, ob er erreichbar war oder auch unvorhergesehene Eventualitäten wie Flugverspätungen zu Gunsten des Gastes abgefedert werden konnten. (Airbnb Ireland, private unlimited company 2016b)

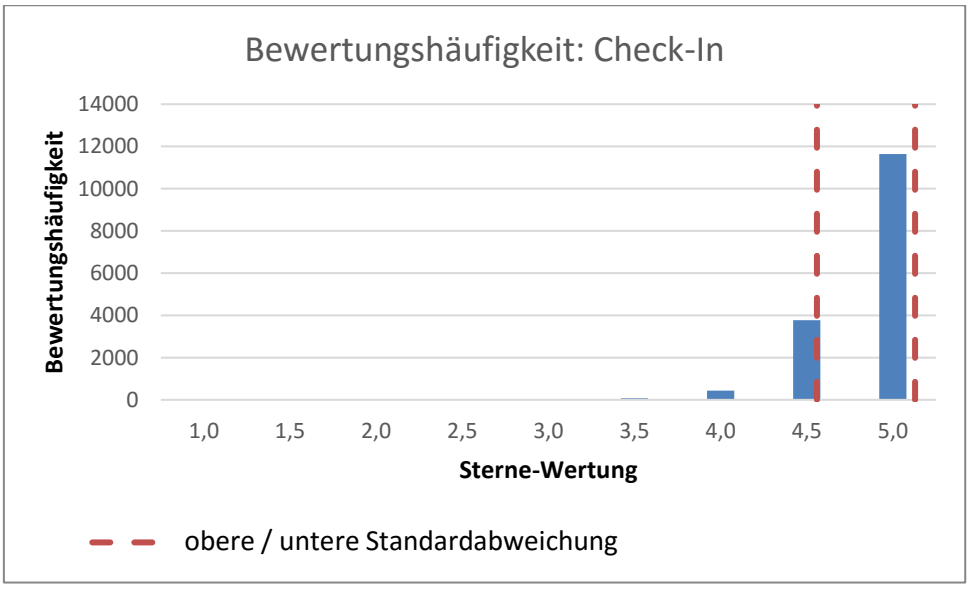


Abbildung 6: Häufigkeitsverteilung Check-In

Als Teil der Gesamtkommunikation (Abbildung 3) differenziert sich das Bild der Häufigkeitsverteilung der Bewertungen des Check-In (Abbildung 6) kaum. Der Mittelwert liegt bei 4,84 und die Standardabweichung ist mit 0,28 am zweitniedrigsten. Damit liegen lediglich Bewertungen von 5,0 Sternen (73 %) innerhalb der Standardabweichung.

Preis/Leistung: Mit dem Preis-Leistungs-Verhältnis wird bewertet, ob der veranschlagte Preis – nach Ermessen des Gastes – gerechtfertigt war. Ähnlich wie bei der Korrektheit der Angaben bezieht sich dieser Bewertungsfaktor auf den gesamten Aufenthalt. (Airbnb Ireland, private unlimited company 2016b)

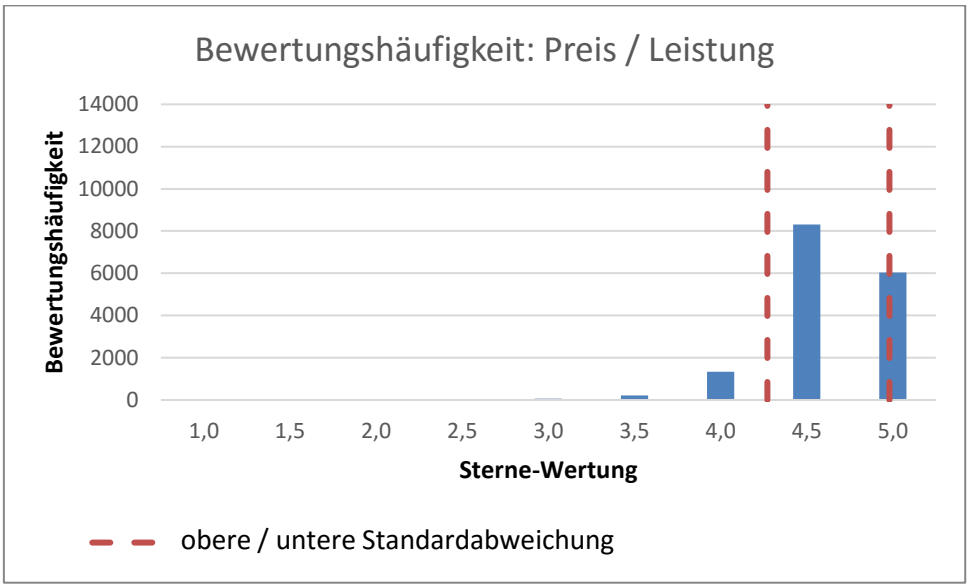


Abbildung 7: Häufigkeitsverteilung Preis/Leistung

Abbildung 7 zeigt, dass bei der Bewertung des Preis-Leistungs-Verhältnis das Maximum bei 4,5 Sternen liegt, während es bei allen vorherigen Kriterien bei 5,0 liegt. Gleichzeitig ist der Mittelwert verglichen mit den übrigen Faktoren mit 4,63 am niedrigsten. Die Standardabweichung liegt bei 0,35, so dass die 5,0 Sterne-Bewertungen wenn auch nur knapp oberhalb der Standardabweichung liegen.

Korrelationen der Bewertungskriterien

Korrelationen

	Kor-rektheit	Kommu-nikation	Sauberkeit	Lage	Check-In	Preis/Leis tung
Korrektheit	1	,454**	,588**	,325**	,467**	,605**
Kommunikation	,454**	1	,414**	,274**	,625**	,473**
Sauberkeit	,588**	,414**	1	,263**	,443**	,564**
Lage	,325**	,274**	,263**	1	,266**	,377**
Check-In	,467**	,625**	,443**	,266**	1	,477**
Preis/Leistung	,605**	,473**	,564**	,377**	,477**	1
Summe	3,440	3,240	3,272	2,505	3,278	3,496

**. Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

Korrelationskoeffizient: Pearson

Werte standardisiert im Bereich –1;+1

Tabelle 1: Korrelationen einzelner Bewertungskriterien

Neben der Häufigkeitsverteilung der einzelnen Bewertungen, werden zudem die Korrelationen zwischen den Variablen untersucht. Tabelle 1 zeigt die Korrelationskoeffizienten *r* nach Pearson im Bereich zwischen –1 und +1. Alle Koeffizienten sind erwartungsgemäß im positiven Bereich.

Nach (Cohen und Manion 1980) besteht eine geringe Korrelation, wenn $0{,}1 \leq r < 0{,}3$, für $0{,}3 \leq r < 0{,}5$ liegt die Korrelation im Mittelmaß und Werte $0{,}5 \geq r$ sind als hoch anzusehen, wobei das Maß stets auf die Fragestellung zu beziehen

ist. In Tabelle 1 sind in der oberen Dreiecksmatrix die Korrelationskoeffizienten hoher Wertigkeit ($r \geq 0{,}5$) in grün, niedriger Wertigkeit ($r \leq 0{,}3$) in orange eingefärbt. Nach gleichem Schema erfolgt die Abstufung in der Summe der einzelnen Faktoren in der letzten Zeile der Matrix. Der größte Korrelationskoeffizient, welcher auch nach (Cohen und Manion 1980) als hoch einzustufen ist, kann zwischen der Kommunikation und dem Check-In festgestellt werden. In der Summe korrelieren aber die Korrektheit der Angaben und das Preis-Leistungs-Verhältnis am stärksten mit den übrigen Kriterien. Den niedrigsten Wert weist die Lage in allen einzelnen Vergleichen und damit auch in der Gesamtsumme auf.

Diskussion

Bewertungen der Wohnungen sind überwiegend bei fünf Sternen

Die Auswertung zeigt, dass die einzelnen Bewertungskriterien zwischen 86 %–97 % mit 4,5–5,0 Sternen gewertet wurden, wovon die 5,0 Sterne-Wertungen einen Anteil zwischen 38 % und 73 % einnehmen. Dieses Ergebnis liegt in einem ähnlichen Bereich, wie das von (Zervas et al. 2015), welcher den Fokus seiner Untersuchung auf die Gesamtbewertung gelegt hat. Das Ergebnis lässt sich somit zunächst auf zwei Weisen deuten. Entweder werden auf Airbnb lediglich Wohnungen angeboten, für die die hohen Werte gerechtfertigt sind, oder es gibt eine gewisse Anzahl an Wohnungen, die vergleichsweise zu gut bewertet wurden. Zur letzteren Erkenntnis gelangen neben (Zervas et al. 2015) ähnlich durchgeführte Studien wie bspw. von (Hu et al. 2009), welcher die Bewertungen von Büchern, DVDs und Videos auf Amazon untersucht hat. Die aus der Untersuchung resultierende J-Form, die sich durch wenig schlechte Bewertungen, überproportional viele sehr gute Bewertungen und praktisch keine mittelmäßigen Bewertungen auszeichnet, kann mit dem vorliegenden Datensatz nicht bestätigt werden. Zwar liegen die meisten Wertungen bei 4,5–5,0 Sternen, doch die schlechtesten Wertungen von 1,0–3,0 Sternen sind so gut wie nicht vorhanden. In der Literatur existieren weitere Arbeiten, die sich mit diesem verzerrten Bewertungsphänomen auseinandersetzen (Chevalier und Mayzlin 2006; Hu et al. 2009), die gezeigt haben, dass Produkte oft gleich bewertet werden und sich Nutzer an vorherigen Bewertungen orientieren. Aufgrund der überwiegenden 4,5–5,0 Sterne-Wertungen im vorliegenden Datensatz und den gezeigten Korrelationen ist davon auszugehen, dass die Einzelwertungen der Gäste nur wenig von der durchschnittlichen Gesamtwertung einer Wohnung abweichen und sich die Bewertungen von Gästen aneinander orientieren. Aus den Arbeiten von (Li und Hitt 2008) lässt sich des Weiteren der Schluss übertragen, dass sich eine gute Wohnung nicht allein daran erkennen lässt, wie viele Sterne sie auf einer Skala hat, sondern dass zusätzliche Bewertungen bspw. in Form von Freitexten nötig sind, sowie die absolute Bewertungshäufigkeit pro Wohnung. Daher ist eine Sterne-Wertung immer in Relation mit der durchschnittlichen Sterne-Wertung aller übrigen Wohnungen zu setzen, sodass auch der Abweichung vom Mittelwert eine gewisse Aussagekraft zugeordnet werden kann. Für die Diskussion treffen wir die Annahme, dass alle Wertungen innerhalb der Standardabweichung den anzunehmenden Normalfall darstellen, Wertungen, die oberhalb der Standardabweichung liegen, überdurchschnittlich gut sind und Werte die unterhalb der Standardabweichung liegen, überdurchschnittlich schlecht sind. Darüber hinaus betrachten wir die Korrelationen zwischen den einzelnen Faktoren im folgenden Teil der Diskussion.

Die Kommunikation ist essentiell und wird bei allen Wohnungsangeboten sehr positiv bewertet

Die Auswertungen der Häufigkeitsverteilungen belegen, dass bei sehr hohen Mittelwerten mit geringen Standardabweichungen eine ausgeprägte Differenzierung zwischen einzelnen Wohnungen verhindert wird. Dies zeigt sich besonders deutlich an der Bewertung der Kommunikation. 75 % der Wohnungen werden mit 5,0 Sternen bewertet, weitere 21 % mit 4,5 Sternen. Dabei liegen die 4,5 Sterne-Wertungen bereits unterhalb der Standardabweichung und sind damit als nicht empfehlenswert einzuordnen. Folglich ist die Verständigung zwischen Gast und Gastgeber aus Sicht des Gasts nur dann gelungen, wenn er 5,0 Sterne vergibt.

Es ist ferner davon auszugehen, das Airbnb-Kunden den Check-In als Teil des Kommunikationsprozesses verstehen, da der erste persönliche Eindruck voneinander meist beim Check-In selbst entsteht. Gastgeber, die schon im E-Mailverkehr positiv auffallen, können diesen Eindruck beim restlichen Ablauf (z. B. Schlüsselübergabe) bestätigen. Diese Hypothese wird von dem im Vergleich höchsten Korrelationskoeffizienten zwischen allen Bewertungskriterien gestützt (r = 0,625). Aufgrund der sehr hohen Korrelation ist auch die Verteilung der Bewertungshäufigkeit ähnlich. 4,5 Sterne-Wertungen liegen bereits unterhalb der Standardabweichung, wodurch lediglich die 5,0 Sterne-Wertungen (73%) der standardmäßigen Erwartungshaltung entsprechen.

Korrekte Angaben fördern das wahrgenommene Preis-Leistungsverhältnis

In der Summe hingegen korrelieren das Preis-Leistungs-Verhältnis, sowie die Korrektheit der Angaben am stärksten mit den übrigen Variablen (vgl. Abbildung 1). Dies lässt sich ggf. dadurch erklären, dass beide Variablen nicht unabhängig von den übrigen Variablen zu betrachten sind. Beide geben Auskunft über den gesamten Aufenthalt. Beim Preis-Leistungs-Verhältnis resümiert der Gast den Aufenthalt und bewertet, ob die Beschreibung der Wohnung, sowie die vorgefundene Sauberkeit etc. mit dem Preis zu rechtfertigen ist. Ähnlich verhält es sich mit der Korrektheit der Angaben. Mit dieser wird vor allem gezeigt, ob die vorgefundene Wohnung der Erwartungshaltung des Gastes entspricht, welcher seine Erwartung u.a. an Bewertungen, Fotos und weiteren Angaben zuvor festmacht (Airbnb Ireland, private unlimited company 2016c). Folglich korrelieren beide Variablen auch miteinander stark (r = 0,605). Das Preis-Leistungskriterium ist dabei die einzige Variable bei der das Maximum an Bewertungen bei 4,5 Sternen liegt. Gleichzeitig liegen alle Wohnungen mit 5,0 Sterne-Wertungen oberhalb der Standardabweichung und sind damit überdurchschnittlich gut.

Die Wohnungslage ist ein sehr unabhängiges Kriterium

Die Lage ist in der Auswertung die Variable, die in jedem Vergleich am wenigsten mit den übrigen Bewertungsfaktoren korreliert. Gäste, die hohen Wert auf gute Bewertungen legen, sollten das Lagekriterium daher besonders beachten, da es im Vergleich zu den übrigen Variablen das Unabhängigste ist. Der geringste Zusammenhang besteht laut Matrix mit der Sauberkeit (r = 0,263). Die Korrelation ist mit der Korrektheit der Angaben hingegen relativ hoch, am höchsten jedoch mit dem Preis (r = 0,377). Daraus lässt sich folgern, dass Wohnungen in weniger attraktiven Lagen, dennoch eine gute Preisbewertung erhalten können. Kommunikation und Check-In sind wie zuvor beschrieben einander relativ ähnliche Bewertungsfaktoren und weisen beide eine mittlere Korrelation mit der Lage auf. Die Standardabweichung umfasst sowohl, 4,5 als auch 5,0 Sterne.

Angaben über die Sauberkeit sind besonders wichtig für eine gute Bewertung

Die Bewertung der Sauberkeit hat einen hohen Korrelationskoeffizienten mit der Korrektheit der Angaben (r = 0,588) und dem empfundenen Preis-Leistungs-Verhältnis (r = 0,564). Wie bereits zuvor beschrieben hat der Gast eine bestimmte Erwartungshaltung gegenüber der gemieteten Wohnung, welche im Falle der Sauberkeit insbesondere von Fotos und Bewertungen abhängen dürfte. Wird die Erwartung erfüllt, so ist die Wohnung nicht nur sauber, sondern auch die Angabe war korrekt. Die relativ hohe Korrelation zwischen der Sauberkeit einer Wohnung und dem veranschlagten Preis zeigt, dass Gäste die eine Wohnung als nicht sauber genug empfinden, auch das Preis-Leistungs-Verhältnis für nicht optimal halten. Für die Vermieter zeigt sich, dass sich vergleichsweise höhere Preise mit einem hohen Grad an Sauberkeit besser rechtfertigen lassen. Die Standardabweichung ist mit 0,42 unter allen Bewertungskriterien am größten und umfasst die 4,5 und 5,0 Sterne-Wertungen.

Sterne sind Glaubenssache, gerade wenn es kaum Unterschiede gibt

Die Diskussion hat bis zu diesem Punkt ergeben, dass eine Wohnung, die in einem der sechs Einzelkriterien weniger als 4,5 Sterne erhalten hat, als überdurchschnittlich schlecht in diesem Kriterium gilt. Der Standard verlangt sogar bei der Kommunikation und dem Check-In 5,0 Sterne, wohingegen bei Top-Wohnungen lediglich das Preis-Leistungsverhältnis noch steigerbar ist, um sich vom Standard abzugrenzen. Das Gäste die Wohnungen dermaßen positiv bewerten, könnte damit zusammenhängen, dass die Wertigkeit eines Sterns nicht definiert und das Bewertungssystem daher nur bedingt transparent ist. (Hu et al. 2009) hat gezeigt, dass Nutzer dazu tendieren, eher positive Bewertungen zu verfassen. Zu untersuchen wäre darüber hinaus, ob Gäste, die lediglich 4,0 Sterne oder 4,5 Sterne vergeben, den Gastgeber durch den Abzug mehr strafen, als dass sie im positiven Sinne 4,0 von 5,0 Sternen vergeben.

Gut gemeinte Bewertungen können negativ sein, wenn sie vom Durchschnitt abweichen

Damit ist gemeint, dass die Intervallabstände der einzelnen Sterne nicht konstant sind. Entsprechen 4,0 von 5,0 Sternen augenscheinlich einer Zufriedenheit von 80 %, so könnte der Gast eigentlich nur zu bspw. 60 % zufrieden sein. Durch die Vergabe von 4,0 Sternen könnte der Gast Kritik zum Ausdruck zu bringen ohne de facto eine schlechte Bewertung abzugeben, wodurch sich der Gastgeber im Nachhinein nicht beschweren kann. Für eine objektive Auswertung ist dies allerdings nicht hilfreich. Des Weiteren ist es nicht möglich eine Kategorie mit weniger als einem Stern zu bewerten (Airbnb Ireland, private unlimited company 2015). Daraus folgt, dass der Kunde eine 0 %-Zufriedenheit mit einem Stern bewerten muss, anstatt intuitiv mit null Sternen. Auch in den vorliegenden Datensätzen von Airbnb gibt es daher keine Einzelbewertung, die schlechter ist als 1,0 Sterne.

Bewertungsdimensionen sind wichtig, auch wenn sie abhängig voneinander sind

Obwohl sechs einzelne Bewertungskriterien sowie die Gesamtwertung genutzt werden, zeigt die Auswertung, dass von den untersuchten Einzelkriterien vier Faktoren miteinander paarweise stark korrelieren und zwei weitere mit den übrigen Faktoren mittelmäßig bis stark korrelieren. Die reine Anzahl an Bewer-

tungsfaktoren ist damit kein ausreichendes Kriterium um festzustellen, ob eine Wohnung empfehlenswert ist. Vielmehr ist es wichtig, dass die Variablen möglichst unabhängig voneinander sind, um ein breitgefächertes Meinungsbild zu erhalten. Airbnb wirkt dem entgegen, indem weitere Informationen in Form von Freitexten, Bildern und Ausstattungsbeschreibungen bereitgestellt werden. Gleichzeitig muss es ausreichend und klar voneinander abgegrenzte Bewertungsfaktoren geben, auch wenn diese miteinander korrelieren. Bspw. wäre es nicht denkbar, dass die Kriterien Preis-Leistungs-Verhältnis und Sauberkeit aus Gründen hoher Korrelation zu einer Wertung zusammengefasst werden, da damit dem Gast wichtige Informationen vorenthalten werden und die Wertung damit noch weniger transparent wäre.

Zusammenfassung

Mittels SPSS wurden für diese Studie die Daten von einer Million Airbnb-Nutzern mit 17 609 bereitgestellten Wohnungen untersucht. Die Auswertung der Sterne-Wertungen hat ergeben, dass alle Bewertungsfaktoren hohe Durchschnittswerte bei geringen Standardabweichungen haben. Darüber hinaus konnte festgestellt werden, dass viele dieser Faktoren mittelmäßig bis stark miteinander korrelieren. 1 657 Wohnungen (9,4 % der Gesamtsumme) konnten nicht in die Statistiken miteinbezogen werden. Dies entspricht fast jeder zehnten Wohnung, die weniger als drei Bewertungen umfasst.

Die anschließende Diskussion hat ergeben, dass die Bewertung eines Gastes maßgeblich von den vorherigen Bewertungen, als auch der persönlichen Erwartungshaltung abhängt. Die Zusammensetzung einer Bewertung ist somit von vielen subjektiven Faktoren abhängig und nur schwer quantifizierbar. Da zu einer eigenen Meinungsbildung möglichst viele Informationen bereitgestellt werden müssen, steht die Forderung nach einem einfachen Bewertungsmodell, wie der analysierten Sterne-Wertung, in einem Zielkonflikt mit der absolut transparenten Nachvollziehbarkeit der einzelnen Bewertung. Darüber hinaus existiert keine klare Definition darüber, wie viel ein einzelner Stern wert ist. So können bspw. nicht weniger als 1,0 Sterne vergeben werden und auch der Wertungsabstand zwischen den einzelnen Sternen muss nicht als konstant angenommen werden. In Folge dessen können Bewertungen nur eine sinnvolle Orientierungshilfe geben, der persönliche Eindruck hängt von den eigenen Erwartungen ab und wie der Gast die Situation vor Ort erlebt.

Die aufgezeigte Problematik gilt allerdings nicht nur für Airbnb, sondern generell für Portale, bei denen Nutzerwertungen einen Einfluss auf die Aktivität auf einer Plattform haben. Das Interesse der Betreiber an Aktivität in Form von vielen Bestellungen oder Buchungen, ausgelöst durch möglichst positive Kommentare, führt zu einer Gradwanderung zwischen Objektivität und Umsatzmaximierung.

Quellverzeichnis

Airbnb Ireland, private unlimited company (2015): Airbnb – Airbnb Community Support. Online verfügbar unter https://airbnb.directly.com/p/airbnb-airbnb-community-support-898398?r=1102577, zuletzt geprüft am 19.06.2016.

Airbnb Ireland, private unlimited company (2016): Kann ich eine Bewertung verfassen, nachdem der 14-tägige Bewertungszeitraum vorbei ist? | Airbnb-Hilfecenter. Online verfügbar unter https://www.airbnb.de/help/article/995/can-i-write-a-review-after-the-14-day-review-period-has-ended, zuletzt geprüft am 19.06.2016.

Airbnb Ireland, private unlimited company (2016): Über uns - Airbnb. Online verfügbar unter https://www.airbnb.de/about/about-us, zuletzt geprüft am 19.06.2016.

Airbnb Ireland, private unlimited company (2016): Wie funktionieren die Sterne-Bewertungen? | Airbnb-Hilfecenter. Online verfügbar unter https://www.airbnb.de/help/article/1257/how-do-star-ratings-work, zuletzt geprüft am 19.06.2016.

Airbnb Ireland, private unlimited company (2016): Wie suche ich nach Unterkünften? | Airbnb-Hilfecenter. Online verfügbar unter https://www.airbnb.de/help/article/252/how-do-i-search-for-listings, zuletzt geprüft am 19.06.2016.

Bolton, Gary; Greiner, Ben; Ockenfels, Axel (2012): Engineering Trust: Reciprocity in the Production of Reputation Information. In: *Management Science* 2012 (59), S. 265–285. Online verfügbar unter http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1120.1609, zuletzt geprüft am 19.06.2016.

Chevalier, Judith A.; Mayzlin, Dina (2006): The Effect of Word of Mouth on Sales: Online Book Reviews. In: *Journal of Marketing Research* 2006 (43), S. 345–354. Online verfügbar unter https://msbfile03.usc.edu/digitalmeasures/mayzlin/intellcont/chevalier_mayzlin06-1.pdf, zuletzt geprüft am 19.06.2016.

Cohen, Louis; Manion, Lawrence (1980): Research methods in education. London: Croom Helm.

Dellarocas, Chrysanthos; Wood, Charles A. (2008): The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias. In: *Management Science* 2008 (54), S. 460–476. Online verfügbar unter http://dl.acm.org/citation.cfm?id=1527788, zuletzt geprüft am 19.06.2016.

Ho, Erica (2015): Why You Really Can't Trust Airbnb Reviews At All. Hg. v. maphappy.org. Online verfügbar unter http://maphappy.org/2015/05/why-you-really-cant-trust-airbnb-reviews-at-all/, zuletzt aktualisiert am 14.05.2015, zuletzt geprüft am 19.06.2016.

Hu, Nan; Zhang, Jie; Pavlou, Paul A. (2009): Overcoming the J-shaped distribution of product reviews. In: *Communications of the ACM - A View of Parallel Computing* 2009 (52), S. 144–147. Online verfügbar unter http://dl.acm.org/citation.cfm?id=1562800.

Li, Xinxin; Hitt, Lorin M. (2008): Self-Selection and Information Role of Online Product Reviews. In: *Information Systems Research* 2008 (19), S. 456–474. Online verfügbar unter http://dx.doi.org/10.1287/isre.1070.0154, zuletzt geprüft am 19.06.2016.

Luca, Michael (2011): Reviews, Reputation, and Revenue: The Case of Yelp.com. Harvard Business School. Online verfügbar unter http://hbswk.hbs.edu/item/reviews-reputation-and-revenue-the-case-of-yelp-com, zuletzt geprüft am 19.06.2016.

Luca, Michael; Zervas, Georgios (2015): Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud. Harvard Business School. Negotiations, Organizations & Markets Unit. Online verfügbar unter http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2293164, zuletzt geprüft am 19.06.2016.

Mayzlin, Dina; Dover, Yaniv; Chevalier, Judith (2014): Promotional Reviews: An Empirical Investigation of Online Review Manipulation. In: *American Economic Review* 2014 (104), S. 2421–2455. Online verfügbar unter https://www.aeaweb.org/articles?id=10.1257/aer.104.8.2421, zuletzt geprüft am 19.06.2016.

Muchnik, Lev; Aral, Sinan; Taylor, Sean J. (2013): Social Influence Bias: A Randomized Experiment. In: *Science* 2013 (341), S. 647–651. Online verfügbar unter http://science.sciencemag.org/content/341/6146/647, zuletzt geprüft am 19.06.2016.

Salganik, Matthew J.; Dodds, Peter S.; Watts, Duncan J. (2006): Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. In: *Science* 2006 (311), S. 854–856. Online verfügbar unter http://science.sciencemag.org/content/suppl/2006/02/06/311.5762.854.DC1, zuletzt geprüft am 19.06.2016.

Stremmel, Jan (2015): Warum bei Airbnb alles wunderbar ist. Hg. v. Süddeutsche Zeitung. Online verfügbar unter http://www.sueddeutsche.de/reise/kolumne-schoen-doof-warum-bei-airbnb-alles-wunderbar-ist-1.2651207, zuletzt geprüft am 19.06.2016.

Worthmann, Merten (2015): Noch warm? Hg. v. Zeit Online. Zeit Online. Online verfügbar unter http://www.zeit.de/2015/48/airbnb-geschaeftsidee-schulung-gastgeber-kritik, zuletzt aktualisiert am 13.12.2015, zuletzt geprüft am 19.06.2016.

Zervas, Georgios; Proserpio, Davide; Byers, John (2015): A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average. Boston University, Boston. Questrom School of Business. Online verfügbar unter http://ssrn.com/abstract=2554500, zuletzt geprüft am 19.06.2016.

Bisher erschienene Bände der Schriftenreihe

Research Papers in Information Systems Management

ISSN 2191-639X (online)
ISSN 2196-8187 (print)

Band 01

Zarnekow, Rüdiger; Kolbe, Lutz M.; Erek, Koray; Schmidt, Nils-Holger

Studie: Nachhaltigkeit und Green IT in IT-Organisationen

ISBN (online) 978-3-7983-2263-9
Published online 2010

Band 02

Repschläger, Jonas; Zarnekow, Rüdiger

Studie: Cloud Computing in der IKT-Branche

ISBN (online) 978-3-7983-2305-6
Published online 2011

Band 03

Zarnekow, Rüdiger; Erek, Koray; Löser, Fabian; Wilkens, Marc

Referenzmodell für ein Nachhaltiges Informationsmanagement

ISBN (online) 978-3-7983-2378-0
ISBN (print) 978-3-7983-2385-8 Preis: 4,90 Euro
Published 2011

Band 04

Erek, Koray; Schmidt, Nils-Holger; Löser, Fabian; Samulat, Peter

Nachhaltigkeitsmanagement bei der Axel Springer AG.
Auf dem Weg zu einer Green IT

ISBN (online) 978-3-7983-2400-8
Published online 2012

Band 05

Erek, Koray; Schmidt, Nils-Holger; Schilling, Thomas

Green IT bei Bayer Business Services

ISBN (online) 978-3-7983-2401-5
Published online 2012

Band 06

Erek, Koray; Schmidt, Nils-Holger; Glau, Thomas

Green IT im IT-Dienstleistungszentrum Berlin

ISBN (online) 978-3-7983-2402-2
Published online 2012

Band 07

Erek, Koray ; Schmidt, Nils-Holger ; Löser, Fabian

Nachhaltigkeitsorientiertes IT-Management
bei einem internen IT-Dienstleister

ISBN (online) 978-3-7983-2403-9
Published online 2012

Band 08

Opitz, Nicky; Erek, Koray; Henseler, Rainer

Green IT im Bundesverwaltungsamt

ISBN (online) 978-3-7983-2485-5
Published online 2012

Band 09

Schmidt, Nils-Holger; Erek, Koray; Kusiak, Katja; Stelzer, Timo

Green IT bei der SAP AG

ISBN (online) 978-3-7983-2486-2
Published online 2012

Band 10

Schmidt, Nils-Holger; Erek, Koray; Kusiak, Katja

Green IT bei der Üstra Hannoversche Verkehrsbetriebe

ISBN (online) 978-3-7983-2487-9
Published online 2012

Band 11

Opitz, Nicky; Erek, Koray; Rekers, Jan; Dahlem, Markus

Green IT bei der Deutschen Bank AG

ISBN (online) 978-3-7983-2488-6
Published online 2012

Band 12

Repschläger, Jonas; Hahn, Christopher; Zarnekow, Rüdiger

Studie: Handlungsfelder im Cloud Computing

ISBN (online) 978-3-7983-2491-6
Published online 2012

Band 13

Repschläger, Jonas; Zarnekow, Rüdiger

Umfrage zur Anbieterauswahl & Markttransparenz im Cloud Computing

ISBN (online) 978-3-7983-2501-2
Published online 2013

Band 14

Limbach, Felix; Kübel, Hannes; Zarnekow, Rüdiger

**Kooperativer Breitbandausbau in Deutschland
Eine Expertenbefragung unter Unternehmensführern und Kooperationsverantwortlichen der deutschen Telekommunikationsbranche**

ISBN (online) 978-3-7983-2590-6
ISBN (print) 978-3-7983-2589-0 Preis: 5,90 Euro
Published 2013

Band 15

Repschläger, Jonas; Zarnekow, Rüdiger; Meinhardt, Nils; Röder, Christoph; Pröhl, Thorsten

**Vertrauen in der Share Economy
Studie: Analyse von Vertrauensfaktoren für Online-Profile**

ISBN (online) 978-3-7983-2775-7
Published online 2015

Band 16

Zarnekow, Rüdiger; Pröhl, Thorsten

Preisvorteile durch frei konfigurierbare Instanzen im Rahmen des Cloud Computing

ISBN (online) 978-3-7983-2839-6
Published online 2016

Universitätsverlag der TU Berlin

Analyse der Wohnungsbewertungen von Airbnb

Mit Millionen von Nutzern bietet Airbnb im Rahmen der Share Economy die Möglichkeit Unterkünfte von Privat zu Privat zu vermieten. Entscheidend für das Zusammenfinden von Mieter und Vermieter ist das gegenseitige Vertrauen, welches auf Airbnb durch gegenseitige Bewertungen der Nutzer und ihrer Wohnungen, sowie Verifikationsnachweisen zustande kommen soll.
Um die Frage zu klären, ob es auf Airbnb überhaupt schlecht bewertete Wohnungen gibt und welchen Einfluss die Bewertungen auf das Nutzerverhalten haben, führen wir eine statistische Auswertung mittels der Daten von einer Million Airbnb-Nutzern durch. Wir evaluieren die Häufigkeitsverteilungen einzelner Bewertungsfaktoren und überprüfen diese auf mögliche Korrelationen.
Mit Hilfe der Ergebnisse wird im Anschluss erklärt, warum vermeintlich gute Bewertungen unterdurchschnittlich sein können, ein hohes Maß an Sauberkeit für einen erfolgreichen Vermieter wichtig ist und auf welche Faktoren Gäste bei der Auswahl ihrer Unterkunft besonders achten sollten, um gute von schlechten Wohnungen zu unterscheiden.

ISBN 978-3-7983-2844-0 (online)

ISBN 978-3-7983-2844-0

http://verlag.tu-berlin.de