STIVEN CATALYST

DRITTE SERIE · NOVEMBER 2026

MANAGEMENT REVIEW

Management ohne Theater.



KPIS

ONLINE-BEWERTUNGEN: WAS EIN STERN WERT IST

Ein halber Stern in Seattle, was ein Punkt einem Hotel wert ist, drei Dinge, die der Durchschnitt verbirgt, das J, das niemand wählte, die Fälschungen und eine Karte für die Bewertungen des Monats.

11,2 %

höheren Preis konnte ein Hotel bei gleicher Buchungschance verlangen, wenn seine Bewertung einen Punkt von fünf höher lag. Travelocity, 2012.

ANDERSON, 2012

IN DIESER AUSGABE

TITELTHEMA

Ein halber Stern in Seattle

SEITE 03

WAS DIE FORSCHUNG SAGT

Das J, das niemand wählte

SEITE 06

WERKZEUG DER AUSGABE

Die Monatskarte der Bewertungen

SEITE 08

NR. 108 · KPIS

In dieser Ausgabe

Sterne sind eine Kennzahl geworden, die der Kunde schreibt und der ganze Markt liest. Eine Führungskraft wird gefragt, warum der Wert von 4,4 auf 4,2 fiel, und fragt selten, wie viele schrieben, wer sie waren oder ob alle echt waren. Diese Ausgabe gibt einem Stern einen Preis und nimmt dann den Durchschnitt auseinander.

In Seattle brachte ein Stern mehr auf Yelp einem unabhängigen Restaurant 5 bis 9 % mehr Umsatz, einer Kette nichts Messbares. Für Hotels auf Travelocity war 2012 ein Punkt mehr 11,2 % höheren Preis wert. Doch auf Amazon sind die meisten Bewertungen Fünfen, mit einem kleineren Haufen Einsen; als eine ganze Gruppe eine CD bewertete, lagen die Werte niedriger und in der Mitte. Auf Yelp wurden rund 16 % der Bewertungen Bostoner Restaurants als verdächtig gefiltert. Eine Monatskarte schließt die Ausgabe.

Stiven Janaqi HERAUSGEBER · STIVEN CATALYST

03 TITELTHEMA
Ein halber Stern in Seattle

04 DIE ZAHLEN
Was ein Punkt einem Hotel wert ist

05 DAS MODELL
Drei Dinge, die der Durchschnitt verbirgt

06 WAS DIE FORSCHUNG SAGT
Das J, das niemand wählte

07 WIE ES GEMESSEN WIRD
Anzahl, Verteilung und die Fälschungen

08 WERKZEUG DER AUSGABE
Die Monatskarte der Bewertungen

09 QUELLEN
Quellen und Methode

SO LESEN SIE DIESE AUSGABE

Zahl
Jede Zahl hat Quelle und Jahr unten auf ihrer Seite.

Unsere Lesart
Wo der Herausgeber deutet und nicht die Forschung, steht es dabei.

Praxis
Schritte und Karte sind Vorschläge zum Ausprobieren, keine Forschungsergebnisse.

TITELTHEMA

Ein halber Stern in Seattle

Michael Luca von der Harvard Business School verknüpfte Yelp-Bewertungen mit den Umsätzen, die jedes Restaurant in Seattle der Steuerbehörde meldete, 2003 bis 2009: 3.582 Restaurants. Yelp zeigt jeden Durchschnitt auf den nächsten halben Stern gerundet, eine 3,24 erscheint also als drei Sterne, eine 3,25 als dreieinhalb.

5–9 %
mehr Umsatz für ein unabhängiges Restaurant je zusätzlichem Stern, je nach Methode

≈ 0
Effekt bei Kettenrestaurants: nicht signifikant

Beim Vergleich von Restaurants knapp über und unter einer Rundungsgrenze hob der zusätzliche halbe Stern allein den Umsatz; auf einen ganzen Stern hochgerechnet rund 9 %. Die Gäste reagierten auf die gerundete Zahl, die sie sahen, nicht nur auf die Bewertungen dahinter. 2009 deckte Yelp 70 % der Restaurants Seattles ab; die Lokalzeitung hatte rund 5 % besprochen.

UNSERE LESART

Die Rundungsgrenze ist willkürlich, und das Geld folgte ihr. Eine angezeigte Zahl ist eine Entscheidung für den Kunden.

Quelle: Michael Luca, Harvard Business School, 2016

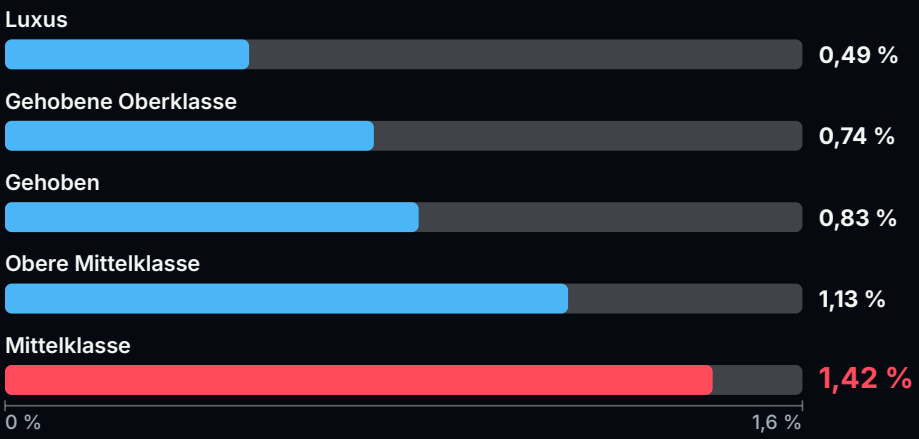
Luca (2016), HBS-Arbeitspapier: Quartalsumsätze der Steuerbehörde des Bundesstaats Washington; 5,4 % aus Fixed Effects, rund 9 % aus Restaurants innerhalb von 0,1 Stern um eine Schwelle. Eine Stadt, Daten bis 2009; Ketten waren rund 5 % der Restaurants.

DIE ZAHLEN

Was ein Punkt einem Hotel wert ist

Chris Anderson von der School of Hotel Administration der Cornell University untersuchte 13.341 Travelocity-Buchungen in neun US-Städten im Juli 2012. Ein Punkt mehr auf der Fünf-Punkte-Bewertung erhöhte die Chance eines Hotels, gewählt zu werden, um 14,2 %: genug, um 11,2 % mehr zu verlangen und gleich oft gebucht zu werden. Jeder Platz tiefer in der Liste senkte die Chance um 11,5 %.

REVPAR-ÄNDERUNG JE 1 % BESSERER ONLINE-REPUTATION, NACH HOTELKLASSE, 2010–2012



UNSERE LESART

Je weniger ein Gast aus Marke und Preis ablesen kann, desto mehr entscheiden die Bewertungen. In der Mittelklasse war der Gewinn fast dreimal so groß wie im Luxussegment.

Quelle: Chris K. Anderson, Cornell University, Center for Hospitality Research, 2012

Anderson (2012): STR-Monatsdaten, Januar 2010 bis Juni 2012, 11 Städte in Europa und Nordamerika, gegen das Wettbewerbsumfeld jedes Hotels; Reputation ist der Index von ReviewPro. Über alle Klassen: Preis 0,80 %, Auslastung 0,20 %, RevPAR 0,96 %. Zusammenhänge, alte Daten; vieles sei nicht gemessen, schreibt er.

DAS MODELL

Drei Dinge, die der Durchschnitt verbirgt

Ein Sternedurchschnitt ist eine Zahl aus drei Dingen, die der Leser nicht sieht: wie viele schrieben, wie ihre Bewertungen verteilt sind und wer sie waren. Luca fand, dass Leser das Erste und das Letzte gewichten, wenn sie es sehen.

01

Wie viele

Eine Änderung wirkte ab 50 Bewertungen 50 % stärker als bei unter 10.

02

Die Verteilung

Bewertungen häufen sich bei Fünf und Eins; der Mittelwert liegt zwischen zwei Lagern.

03

Wer schrieb

Yelps Elite-Rezensenten wirkten etwa doppelt; manche Rezensenten sind keine Kunden.

BEISPIEL, VON DER REDAKTION ERFUNDEN

Hotel A, 10 Bewertungen	alle drei Sterne: Durchschnitt 3,0
Hotel B, 10 Bewertungen	fünf Fünfen, fünf Einsen: Durchschnitt 3,0

Gleicher Durchschnitt, zwei verschiedene Hotels: eines ist für alle gewöhnlich, das andere spaltet seine Gäste. Nur die Verteilung trennt sie.

Quellen: Michael Luca, Harvard Business School, 2016; Nan Hu, Paul A. Pavlou & Jennifer Zhang, 2007; Michael Luca & Georgios Zervas, Management Science 62(12), 2016

Karten: Zusammenfassung der Redaktion aus Luca (2016, Tabellen 9 und 10), Hu, Pavlou & Zhang (2007) und Luca & Zervas (2016). Die 50 % vergleichen Restaurants mit mindestens 50 Bewertungen mit solchen unter 10. Das Beispiel ist erfunden.

WAS DIE FORSCHUNG SAGT

Das J, das niemand wählte

Nan Hu, Paul Pavlou und Jennifer Zhang sammelten 2005 Amazon-Bewertungen zu 32.878 Büchern, 17.978 DVDs und 28.983 Videos. Die Bewertungen häuften sich bei fünf Sternen, mit einem kleineren Haufen bei einem und wenig dazwischen: ein J. Dann gaben sie 66 Studierenden eine Musik-CD und baten jeden, sie zu bewerten.

ANTEIL DER AMAZON-BEWERTUNGEN MIT VIER ODER FÜNF STERNEN, 2005



60 Studierende antworteten. Ihre Bewertungen bildeten eine Glocke um die Mitte, Mittelwert 3,12; auf Amazon hatte dieselbe CD im Mittel 3,54, mit über 60 % der Bewertungen bei einem oder fünf Sternen, gegen unter 10 % bei den Studierenden. Die meisten Produkte hatten zudem wenige Stimmen: 78 % der Bücher hatten weniger als 20 Bewertungen.

UNSERE LESART

Wer kauft, mag das Produkt schon, und wer schreibt, ist begeistert oder verärgert. Der gewöhnliche Kunde schweigt meist.

Quelle: Nan Hu, Paul A. Pavlou & Jennifer Zhang, 2007

Hu, Pavlou & Zhang (2007), Arbeitspapier; Amazon-Daten Februar bis Juli 2005; Balken sind Anteile an allen Bewertungen der Kategorie. Die Autoren nennen die zwei Ursachen Kaufverzerrung und Untermeldungsverzerrung. Eine CD und Studierende: Grenzen, die sie selbst nennen.

WIE ES GEMESSEN WIRD

Anzahl, Verteilung und die Fälschungen

Michael Luca und Georgios Zervas untersuchten 316.415 Yelp-Bewertungen zu 3.625 Bostoner Restaurants, 2004 bis 2012. Yelps Filter hatte rund 16 % als verdächtig aussortiert. Gefilterte Bewertungen waren extremer; Restaurants mit wenigen oder frisch schlechten Bewertungen betrogen eher, Ketten seltener.

79 % : 19 %

gefilterte Bewertungen bei Betrieben, die beim Anwerben falscher Bewertungen ertappt wurden, gegen das durchschnittliche Bostoner Restaurant

01	**Anzahl**	wie viele diesen Monat, und die Gesamtzahl hinter dem Durchschnitt
02	**Verteilung**	der Anteil der Einsen und Zweien und der Fünfen, nicht nur der Mittelwert
03	**Aktuell, nicht gesamt**	der Monat gegen die letzten drei; der Gesamtdurchschnitt bewegt sich langsam
04	**Wer schrieb**	gefilterte Bostoner Rezensenten hatten im Mittel 10 Bewertungen geschrieben, veröffentlichte 137

Quelle: Michael Luca & Georgios Zervas, Management Science 62(12), 2016

Luca & Zervas (2016), angenommene Fassung von 2015. Der Filter ist ein Näherungswert für Fälschungen, mit Fehlern in beide Richtungen; die Testaktion umfasste 126 US-Betriebe, die Yelp in den 90 Tagen bis März 2014 markierte. Die Felder sind ein Vorschlag der Redaktion. Top-2-Box steht in Nr. 32, Goodharts Gesetz in Nr. 43.

WERKZEUG DER AUSGABE

Die Monatskarte der Bewertungen

Einmal im Monat die Bewertungen als Beschwerdebuch lesen, nicht als Wert: zählen, die Verteilung ansehen, die schlechten nach Ursache ordnen und prüfen, ob die Rezeption es zuerst hörte. Das Werkzeug Guest Complaints macht aus Kategorien ein Pareto und findet die Zimmer, die zweimal klagen.

01	**Monat, Plattform, Anzahl**	neue Bewertungen diesen Monat; die Gesamtzahl hinter dem Durchschnitt
02	**Durchschnitt und Verteilung**	Anteil Einsen und Zweien, Anteil Fünfen; gegen die letzten drei Monate
03	**Was die schlechten sagen**	in den Worten des Gastes, je eine Kategorie
04	**Zuerst an der Rezeption gehört?**	im Beschwerdebuch oder nicht; wenn nicht, warum
05	**Behoben und beantwortet**	wer die Ursache behebt, bis wann; wer die Bewertung beantwortet

Quellen: Michael Luca, Harvard Business School, 2016; Nan Hu, Paul A. Pavlou & Jennifer Zhang, 2007; Michael Luca & Georgios Zervas, Management Science 62(12), 2016

Die Karte ist ein Vorschlag der Redaktion, keine erprobte Methode, gestützt auf die Warnungen der drei Quellen zu Anzahl, Verteilung und Verfassern. Nach Ansicht der Redaktion sollte ein Bewertungswert kein Ziel für das Personal werden. Der unten verlinkte Essay zeigt, wie der Autor eine Beschwerde in den Worten des Gastes festhält.

Mehr im Essay Der verärgerte Gast um 2 Uhr nachts

Tool öffnen: Gästebeschwerden stivencatalyst.com/de/tools/guest-complaints/

QUELLEN UND METHODE

Jede Zahl hat eine Quelle.

Die Zahlen dieser Ausgabe stammen aus den folgenden Quellen. Das Jahr zeigt, wie aktuell jede ist.

Reviews, Reputation, and Revenue: The Case of Yelp.com (Working Paper 12-016)
Michael Luca, Harvard Business School, 2016
https://www.hbs.edu/ris/Publication%20Files/12-016_a7e4a5a2-03f9-490d-b093-8f951238dba2.pdf

The Impact of Social Media on Lodging Performance (Cornell Hospitality Report 12(15))
Chris K. Anderson, Cornell University, Center for Hospitality Research, 2012
https://ecommons.cornell.edu/items/d64bbbba-651a-4e13-a058-af3e2e82f4cf

Why do Online Product Reviews have a J-shaped Distribution? Overcoming Biases in Online Word-of-Mouth Communication (working paper)
Nan Hu, Paul A. Pavlou & Jennifer Zhang, 2007
https://archive.nyu.edu/bitstream/2451/14951/2/usedbook1.pdf

Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud
Michael Luca & Georgios Zervas, Management Science 62(12), 2016
https://doi.org/10.1287/mnsc.2015.2304

REDAKTIONELLE METHODE

Jede Zahl wurde auf Jahr, Herausgeber und Bedeutung geprüft, am Volltext, wo er sich öffnen ließ. Eine nur in einer Zusammenfassung, einem Abstract oder Auszügen gelesene Quelle ist mit „über“ markiert. Eine Zahl, die sich im Gelesenen nicht bestätigen ließ, ist auf ihrer Seite so markiert oder fehlt. Die Deutung des Herausgebers ist als „Unsere Lesart“ markiert.



MANAGEMENT REVIEW

Management ohne Theater.

Jede Ausgabe eine Führungsfrage, geprüft an der besten Forschung.

ALLE AUSGABEN

stivencatalyst.com/de/magazine/management-review.html

Management Review · Nr. 108 · November 2026 · Stiven Catalyst
Herausgegeben von Stiven Janaqi, der sie verantwortet; „die Redaktion" ist in dieser Publikation er. Texte, Übersetzungen, Layout und die Prüfung der Zahlen an ihren Quellen entstehen in Arbeitssitzungen mit Claude, einem KI-Modell von Anthropic.
stivencatalyst.com/de/how-this-is-made.html