

---

**essentials**

*essentials* liefern aktuelles Wissen in konzentrierter Form. Die Essenz dessen, worauf es als „State-of-the-Art“ in der gegenwärtigen Fachdiskussion oder in der Praxis ankommt. *essentials* informieren schnell, unkompliziert und verständlich

- als Einführung in ein aktuelles Thema aus Ihrem Fachgebiet
- als Einstieg in ein für Sie noch unbekanntes Themenfeld
- als Einblick, um zum Thema mitreden zu können

Die Bücher in elektronischer und gedruckter Form bringen das Expertenwissen von Springer-Fachautoren kompakt zur Darstellung. Sie sind besonders für die Nutzung als eBook auf Tablet-PCs, eBook-Readern und Smartphones geeignet. *essentials*: Wissensbausteine aus den Wirtschafts-, Sozial- und Geisteswissenschaften, aus Technik und Naturwissenschaften sowie aus Medizin, Psychologie und Gesundheitsberufen. Von renommierten Autoren aller Springer-Verlagsmarken.

Weitere Bände in der Reihe <http://www.springer.com/series/13088>

---

Paul Niebler · Dominic Lindner

# Datenbasiert entscheiden

Ein Leitfaden für Unternehmer und  
Entscheider

 Springer Gabler

Paul Niebler  
IBM Deutschland GmbH  
Frankfurt am Main, Deutschland

Dominic Lindner  
noris network AG  
Nürnberg, Deutschland

ISSN 2197-6708

ISSN 2197-6716 (electronic)

essentials

ISBN 978-3-658-23927-5

ISBN 978-3-658-23928-2 (eBook)

<https://doi.org/10.1007/978-3-658-23928-2>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Gabler

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2019

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Springer Gabler ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

---

## Was Sie in diesem *essential* finden können

- Entscheidungen auf der Grundlage von Daten treffen
- Tipps für Unternehmer zur effektiven Nutzung von Daten
- Statistische Grundlagen für ein besseres Verständnis von Daten
- Abgrenzung der Buzzwords im Bereich der Data Science
- Potenziale und Risiken von Daten
- Praktisches Beispiel zur Nutzung und Auswertung von Daten

---

## Vorwort

„Du kannst Daten ohne Informationen haben, aber du kannst nicht Informationen ohne Daten erhalten“ (Daniel Keys Moran). So steigt die Anzahl an gespeicherten Daten täglich. Bereits heute werden jeden Tag geschätzt 2,5 Trillionen Daten (WiWo) erzeugt.

Der Hype um Daten wird von Tag zu Tag größer und auch wir erleben dies in unserer täglichen Arbeit (Lindner et al. 2018). Wir, das sind die Autoren dieses *essentials*. Die Idee dazu kam uns nach einem fachlichen Austausch zum Thema Daten, in dem wir feststellten, dass wir zwar aus verschiedenen Branchen kommen, allerdings das gleiche Ziel und die gleiche Herausforderung haben: Uns steht eine große Menge an Daten zur Verfügung, die aktuell zielgerichtet ausgewertet werden soll. Die Auswertung von Daten führt zu einem besseren Verständnis des eigenen Unternehmens – während früher viele Entscheidungen aus dem Bauch heraus getroffen wurden, sind heute die resultierenden Erfolge durch das gestiegene Datenvolumen so gut messbar wie nie zuvor. Im Optimalfall führt das zu besseren Entscheidungen, die Praxis zeigt uns aber, dass ein datenbasiertes Vorgehen in der Praxis oftmals alles andere als trivial ist. Eine geringe Datenqualität und fehlendes Know-how zu den Grundlagen der Statistik und den am Markt verfügbaren Softwarelösungen stehen der Auswertung von vorhandenen Daten oftmals im Weg.

Paul Niebler ist IT-Berater mit einem Schwerpunkt auf Data Science und unterstützt Kunden hauptsächlich in der Automobilbranche. Sein Ziel ist es, aus den bestehenden Daten neues Wissen zu extrahieren und diesen Erkenntnisgewinn verständlich an Fachabteilungen und Entscheidungsträger zu kommunizieren. Dominic Lindner ist Associate Manager beim Rechenzentrumsprovider noris network AG und promoviert im Bereich der Mittelstandsforschung an der FAU Erlangen-Nürnberg. Vor allem die Speicherung und die schnelle Auswertung aus

Sicht der Infrastruktur sind eine wichtige Herausforderung in seiner täglichen Arbeit. In diesem *essential* liefert er spannende Einblicke aus der wissenschaftlichen Perspektive auf das Thema sowie praktische Tipps aus der Sicht eines Entscheiders in einem KMU.

Wir wollen mit diesem *essential* Entscheidern aus Mittelstand und Konzernen einen Einblick in unsere Arbeit sowie einen praktischen Guide für die Herbeiführung von Entscheidungen auf Basis von Daten geben.

Paul Niebler  
Dominic Lindner

---

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b> .....	1
<b>2</b>	<b>Abgrenzung der Begriffe im Bereich Data Science</b> .....	7
2.1	Data Science .....	7
2.2	Big Data .....	9
2.3	Data Mining .....	10
2.4	Machine Learning .....	10
<b>3</b>	<b>Vorbereitung der Daten im Unternehmen</b> .....	13
3.1	Datenspeicherung .....	13
3.2	Datenverarbeitung .....	14
3.3	Datentypen .....	16
<b>4</b>	<b>Daten im Unternehmen zielführend auswerten</b> .....	19
4.1	Tabellenkalkulationsprogramme .....	19
4.2	Datenvisualisierung .....	20
4.3	Data Mining und Machine Learning .....	21
4.4	Programmiersprachen .....	22
<b>5</b>	<b>Praxisbeispiel</b> .....	23
5.1	Einleitung .....	23
5.2	Geschäftssituation .....	25
5.3	Verfügbare Daten .....	25



---

5.4	Optimierung der Bestellungen.....	27
5.5	Optimierung der Mitarbeiterauslastung.....	29
5.6	Zusammenfassung.....	33
<b>6</b>	<b>Fazit und Tipps für Unternehmer.....</b>	<b>35</b>
	<b>Literatur.....</b>	<b>41</b>