

Was ist DATA SCIENCE

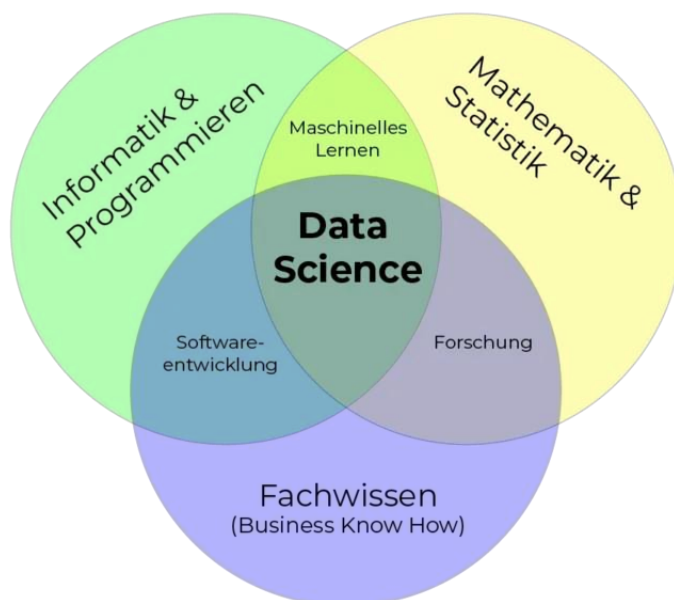
Quelle: <https://datasolut.com/was-ist-data-science/>

Data Science (Datenwissenschaft) ist eine interdisziplinäre Wissenschaft, die sich mit der **Gewinnung von Wissen aus Daten** befasst. Die Datenwissenschaft generiert Informationen aus großen Datenmengen, um daraus **Handlungsempfehlungen für das Unternehmensmanagement** abzuleiten. Ziel dieser Handlungsempfehlungen ist die **Verbesserung der Qualität unternehmerischer Entscheidungen und der Effizienz von Arbeitsabläufen**.

1. [Data Science Definition – Was ist Data Science?](#)
2. [In welchen Bereichen wird Data Science angewandt?](#)
3. [Was sind die Tätigkeiten eines Data Scientist?](#)
4. [Worauf kann man sich als Data Scientist spezialisieren?](#)

Data Science Definition – Was ist Data Science?

Data Science ist die Schnittmenge zwischen den Wissenschaftsbereichen Mathematik, Informatik sowie dem branchenspezifischen Fachwissen. Der Bereich Data Science befasst sich mit der Analyse von (großen) Datenmengen, der Erkennung von Anomalien in den Daten sowie mit der Vorhersage von zukünftigen Ereignissen. Die im Arbeitsbereich Datenwissenschaft arbeitenden Personen werden als **Datenwissenschaftler oder Data Scientist** bezeichnet



Data Science ist die Schnittmenge zwischen den Wissenschaftsbereichen Mathematik, Informatik sowie dem branchenspezifischen Fachwissen.

Erstmals wurde der Begriff „Data Science“ im **Jahr 1960** durch den dänischen Informatik-Pionier Peter Naur als Synonym für „Informatik“ verwendet. 2001 begründete der US-amerikanische Computerwissenschaftler William S. Cleveland die Datenwissenschaft als eigenständige Fachdisziplin. Cleveland benannte als Fachgebiete von Data-Science: Daten-Modelle und Daten-Methodik, Rechenoperationen mit Daten, die Bewertung von Werkzeugen, Pädagogik, Theorie und multidisziplinäre Untersuchungen.

Heute wird die Definition von Data Science deutlich weiter gefasst als noch vor fünfzig Jahren: die Datenwissenschaft befindet sich im **Schnittpunkt zwischen den Wissenschaftsbereichen Mathematik (insbesondere Stochastik) und Informatik sowie dem branchenspezifischen Fachwissen.**

Die Ziele und der eigentliche Zweck vom Einsatz von Data Science lässt sich auf eine breite Masse an Anwendungen beziehen. Zu den Zielen gehören vor allem:

- Den Datenbestand so zu analysieren, dass wertvolle Informationen ersichtlich werden
- Durch die Analyse des Datenbestands soll die Ableitung von Handlungsempfehlungen geschehen
- Data Science als Instrument zur verbesserten Entscheidungsfindung nutzen
- Erstellung von Vorhersagen
- Automatisierung und Optimierung von jeglichen Prozessen innerhalb des Unternehmens

In diesem Video wird erklärt, was ein Data Science ist und welche Fähigkeiten ein Datenwissenschaftler benötigt. (<https://youtu.be/4rPzdf93fgg>)

In welchen Bereichen wird Data Science angewandt?

Data Science lässt sich in etlichen digitalen Gebieten einsetzen, sodass fast keinerlei Grenzen gesetzt sind. Demnach spielt Data Science überall dort eine entscheidende Rolle, in denen eine Vielzahl an Daten aufkommt und Basis dieser Datenmengen unterschiedliche Funktionen wie beispielsweise das Treffen von Prognosen erreicht werden soll. Dennoch gibt es typische Bereiche, in denen Data Science vermehrt vorkommt. Dazu gehören vor allem der Online-Handel bzw. E-Commerce, die Logistik, sowie das Gesundheitswesen, Finanzwesen, Industrie und Produktion. Im Folgenden werden Ihnen die relevantesten Einsatzgebiete aufgeführt.

Neben den Branchen lassen sich die Anwendungen ebenfalls nach Art der Tätigkeit bestimmen. Dabei sind folgende Tätigkeiten zu erwähnen:

- **Explorative Datenanalyse:** Analyse sowie Auswertung vorliegender Daten zur Stützung von Hypothesen.

- **Vorhersage von Wahrscheinlichkeiten:** Kauf- oder Kündigungswahrscheinlichkeiten lassen sich dank Data Science und zugehörigen Verfahren berechnen und vorhersagen
- **Vorhersage von numerischen Werten:** Durch historische Berechnungen ist es möglich, zukünftige Vorhersagen, wie beispielsweise zukünftigen Stromverbrauch oder Umsätze zu treffen.
- **Erkennung von Anomalien:** Data Science bietet die Möglichkeiten Unregelmäßigkeiten und Anomalien in den Daten zu identifizieren.
- **Analyse von Text und Sprache:** Natural Language Processing (NLP) ist eine Möglichkeit, vorhandene Texte sowie gesprochener Sprache auszuwerten
- **Analyse von Bildern und Videos:** Bildererkennung, Klassifikation, usw.
- **Erkennen von Zusammenhängen und Gruppen:** Innerhalb riesiger Datenmengen (Big Data) ist es eine der Aufgaben von Data Science, Zusammenhänge sowie Gruppen innerhalb jeweiliger Datenmengen zu erkennen.

Branchen-Beispiele für den Einsatz von Datenwissenschaften

Große Datenmengen werden heute mit den Instrumenten und Methoden der **Data-Science** für **Unternehmen aller Branchen** ausgewertet.

- **Marketing:** Vor allem zur Personalisierung im Marketing findet Data Science hinsichtlich der Auswertung von Daten große Anwendung. Dabei werden Vorlieben sowie das Verhalten analysiert, sodass sich Handlungsempfehlungen für das Marketing ableiten lassen.
- **(IT)-Security:** Data Science wird zunehmend zur Überwachung von IT-Systemen eingesetzt. Kritische IT-Systeme werden dabei mithilfe von Security Information and Event Management (SIEM) geschützt.
- **Mobilität:** Im Bereich der Mobilität wird autonomes Fahren dank Data Science und maschinellem Lernen weiter vorangetrieben. Dazu werden vorwiegend Sensordaten ausgewertet, um genaue Informationen des Fahrzeugs sowie der Umgebung zu generieren.
- **Retail- und Handelsunternehmen** profitieren von Data-Science durch Analysen des Kaufverhaltens von Kunden. Die Untersuchung möglicher Ursachen für Retouren hilft bei der Verringerung von Warenrücksendungen.
- In der **Gesundheitsbranche (Medizin und Pharmazie)** ermöglicht Data Science die Erstellung von Ähnlichkeitsanalysen als Grundlage für eine individualisierte Behandlung von Patienten und die Optimierung der Medikation.
- **Logistikunternehmen** verbessern mithilfe von Data Science ihre Arbeitsprozesse und die Qualität ihrer Transport-Dienstleistungen.
- **Industriebetriebe** steuern und optimieren Fertigungsabläufe durch den Einsatz von Data Science.
- **Versicherungen und Banken** schöpfen mithilfe von Data-Science das Potenzial der ihnen zur Verfügung stehenden externen und internen Daten aus, um ihre Produkte zu verbessern und die Vertriebsfolge zu steigern.

Wo wird Data Science in der Praxis angewandt?

Im Alltag begegnet uns künstliche Intelligenz schon in unterschiedlichsten Bereichen. Dazu gehören:

- **Virtuelle Assistenten:** Die Benutzung von mobilen Endgeräten wird dank sprachgesteuerten Assistenten wie Alexa (Amazon), Siri (Apple) und Google Assistant Tag für Tag einfacher. Diese Systeme arbeiten größtenteils auf Basis von Natural Language Processing-Systemen, welche menschliche Sprache verstehen und verarbeiten.
- **Chatbots:** Routineaufgaben innerhalb von Kundenserviceabteilungen lassen sich dank Chatbots zum Großteil reduzieren. Diese gewonnene Zeit lässt sich von jeweiligen Mitarbeitern für weitere Aufgaben wie beispielsweise eine individuelle Kundenbetreuung nutzen.
- **Empfehlungssysteme:** Je nach persönlichem Geschmack, gewähren Empfehlungssysteme dem Kunden eine unterschiedliche Auswahl, die den Vorlieben des Kunden entsprechen.
- **Customer Relationship Management:** Künstliche Intelligenz eignet sich im CRM hervorragend, um vorliegendes Marketingbudget effizient an prognostizierten Kundenwerten und Produktaffinitäten zu verteilen
- **Gesichtserkennung:** Im Bereich der Bilderkennung ist es dank künstlicher Intelligenz möglich, innerhalb von sozialen Netzwerken, Gesichter zu erkennen und diese auf Bildern zu markieren.
- **Betrugserkennung:** Im Rahmen der Kriminalitätsbekämpfung wird künstliche Intelligenz derweilen von Finanzdienstleistern, Banken und weitere Unternehmensbranchen zur Betrugserkennung genutzt.
- **E-Commerce:** Erhebliche Kundenbindungsmaßnahmen sowie Umsatzsteigerungen sind dank künstlicher Intelligenz im E-Commerce möglich. Dazu errechnet die künstliche Intelligenz mögliche Produktempfehlungen (Beispiel: Amazon) und schlägt diese potentiellen Kunden vor.
- **Autonomes Fahren:** Das Einhalten des Sicherheitsabstands, Einparken oder Einhalten der Fahrspur ist dank intelligenter Systeme schon derzeit möglich. Doch dies ist erst der Anfang. Vorreiter und Weltkonzern Tesla hat es sich zur Aufgabe gemacht, in naher Zukunft vollautonome Fahrzeuge auf den Markt zu bringen.
- **Supermarkt:** Im Bereich des Lebensmitteleinzelhandels ist Amazon Go eines der bekanntesten Beispiele. Dank Amazon Go soll die Effizienz sowie das Kundenerlebnis während eines Einkaufs drastisch gesteigert werden.

Was sind die Tätigkeiten eines Data Scientist?

Die nationale Forschungsgemeinschaft der US-Regierung wies 2005 auf die **entscheidende Bedeutung der Data Scientists für das Management digitaler Daten** hin. Zu den Datenwissenschaftlern zählen insbesondere Informatiker, Programmierer, Datenbank-Experten, Domain-Spezialisten, Archivare und Bibliothekare und sogar Fachleute der Softwareentwicklung.

Welche Fachkenntnisse braucht ein Data Scientist?

Ein Data Scientist muss über ein breites Spektrum von **Fachkenntnissen** verfügen, um seine analytischen, beratenden und koordinierenden Aufgaben an der Schnittstelle zwischen Datenbanken, Mitarbeitern und Management eines Unternehmens erfüllen zu können:

- **Mathematische und stochastische Modelle und Methoden** bilden die Grundlage für die Analyse und Interpretation von Daten. Von großer Bedeutung ist insbesondere eine gute analytische Fähigkeit, die es ermöglicht Informationen aus Daten zu extrahieren.
- Im Zeitalter großer Datenmengen („Big Data“) ist Datenwissenschaft ohne fundierte Kenntnisse über Datenverarbeitung und ihre Instrumente kaum vorstellbar. **Die Computer-Wissenschaft (Informatik)** unterstützt die Big-Data-Analyse durch Algorithmen, die für eine effiziente Bearbeitung der jeweiligen Datentypen geeignet sind. Analyse-Tools selektieren aus strukturierten oder unstrukturierten Datenmengen die benötigten wertvollen Informationen.
- **Spezifische Branchen-Kenntnisse** sind unerlässlich, um Daten und Geschäftsprozesse eines bestimmten Unternehmens einordnen, analysieren sowie sinnvolle Handlungsempfehlungen entwickeln zu können. **Betriebswirtschaftliches Wissen** ist von grundlegender Bedeutung, wenn Datenanalysen zwecks einer Geschäftsoptimierung durchgeführt werden sollen.

Welche persönlichen Qualifikationen braucht ein Data Scientist?

Das von Data Scientists angewandte Wissen und die dementsprechende Qualifikation haben Data Scientists entweder durch den Abschluss eines datenspezifischen Studiengangs erlangt oder haben sich dieses Wissen eigenständig beigebracht. Bei einer Vielzahl an Data Scientists handelt es sich jedoch oftmals um gelernte Informatiker, Physiker, Statistiker, Betriebswirtschaftler oder Programmierer.

Neben fachlicher Expertise sollte ein Data Scientist in der Lage sein, dass durch die Analyse der Daten generierte Wissen dem Unternehmen oder weiteren Personen verständlich darstellen zu können.

Da ein Data Scientist eine beratende Funktion im Unternehmen einnimmt, sollten die Kommunikations- sowie Präsentationsfähigkeiten auf einem guten Niveau sein.

Die Ausübung von Data Scientist Tätigkeiten setzt zudem einige besonders ausgeprägte **persönliche Qualifikationen** voraus:

- **Sprachliche und kommunikative Fähigkeiten**, um einerseits Branchenexperten zu verstehen und andererseits Mitarbeitern aller Unternehmensbereiche, die aus der Datenanalyse abgeleiteten Handlungsempfehlungen überzeugend vermitteln zu können
- **Kreativität** bei der flexiblen Anpassung analytischer Methoden an neue Herausforderungen und Kontexte
- **Aufgeschlossenheit** bei der Erforschung und Nutzung neuer Analyse-Instrumente und Data-Science-Verfahren sowie

- **Koordinationsvermögen** sowohl hinsichtlich der Delegation von Aufgaben zur Datenbeschaffung als auch bei der Steuerung und Kontrolle von Data-Science-Projekten.

Am Arbeitsmarkt übersteigt derzeit die Nachfrage nach Data Scientists die Anzahl der verfügbaren ausgebildeten Datenwissenschaftler. Zwar liegen weltweit zunehmend umfangreiche Datenbestände vor, doch können diese aufgrund der nicht ausreichenden Verfügbarkeit von Datenwissenschaftlern nicht vollständig analysiert werden.

Worauf kann man sich als Data Scientist spezialisieren?

Nach einer Data Science Ausbildung lassen Sie sich in vielen Bereichen, je nach persönlichen Neigungen und Fähigkeiten, einsetzen. Im Laufe der Zeit und mit der Zunahme an Projekten und Erfahrungen ist es möglich, sich auf spezielle Anwendungsfelder zu spezialisieren. Dabei werden folgende Anwendungsgebiete oftmals bevorzugt:

- **Risikoberatung:** Die Vorhersage der Auswirkungen verschiedener Unternehmensentscheidungen wird im Rahmen der Risikoberatung getroffen. Da dieser Bereich immer wichtiger wird, erweist er sich für viele Data Scientists als attraktive Spezialisierung.
- **Customer Relationship Management:** Im Bereich des Customer Relationship Managements lassen sich unterschiedliche Analysen und Prognosen bezüglich des Kundenverhaltens treffen. Beispielsweise kann ein zuständiger Mitarbeiter analysieren, warum bestimmte Produkte öfters gekauft werden oder aber warum bestimmte Produkte eine hohe Retourenquote aufweisen.
- **Produktpersonalisierung:** Data Scientists können sich ebenfalls auf die Produktspezialisierung fokussieren. In diesem Bereich geht es meist darum, wie gewisse Produkte oder Dienstleistungen persönlicher gestaltet werden können, um dem Kunden bestmöglich zufrieden zu stellen.
- **Marketing:** Ebenso bietet es sich an, seine Fähigkeiten im Bereich des Marketings einzusetzen. Dahingehend lassen sich Kundendaten gezielt analysieren und effizienter nutzen, um dem Kunden personalisierte Angebote anzubieten. Deshalb sind viele Unternehmen auf der Suche nach qualifizierten Data Scientists, die ihre Fähigkeiten im Marketing unter Beweis stellen.
- **Logistik:** Durch die Ansammlung riesiger Datenmengen (Big Data) in Verbindung mit IoT lassen sich in immer mehr Bereichen von Logistikprozessen Daten analysieren. Diese Daten können beispielsweise dazu genutzt werden, die Effizienz von Logistikprozessen, Fahrzeugrouten oder der generellen Dienstleistung zu steigern. Da der gesamte Onlinehandel wächst, profitiert ebenfalls die Logistikbranche. Wer als Data Scientist einen zukunftsicheren Job möchte, ist im Bereich der Logistik mit Sicherheit gut beraten.