



CAREER TRACK // CURRICULUM

# Data Scientist – Fokus Python

Qualifizierung für die Jobrolle als Data Scientist

# Warum dieser Kurs



Data Science ist ein Berufs- und Themenfeld mit unglaublichem Potenzial, das auch in den nächsten Jahren weiter stark wachsen wird. Die Nachfrage nach qualifizierten Data Scientists ist sehr hoch, da in diesem Bereich großer Fachkräftemangel herrscht. Daher gilt die Jobrolle als Data Scientist branchenübergreifend als eine der erstrebenswertesten und nachgefragtesten Top-Karrieren.

Nicht zuletzt werden Data Scientist Weiterbildungen deshalb immer relevanter und nachgefragter. Die Fähigkeiten von Data Scientists sind im Zeitalter der Digitalisierung so wichtig wie begehrt: Fachkompetenzen in der Datenanalyse, Programmierkenntnisse in Python sowie die Anwendung von maschinellem Lernen und Deep Learning.

Die zertifizierte Online-Weiterbildung zum Data Scientist – Fokus Python befähigt Dich dazu, Prognosemodelle aus Daten abzuleiten, zu verifizieren und zu interpretieren, um die Modellergebnisse effizient zu kommunizieren. Der zusätzliche Aufbau von Fähigkeiten im Bereich des Machine Learnings qualifiziert Dich mit erfolgreichem Abschluss des Karrierepfades für die Jobrolle als Data Scientist oder eine andere analytische Jobrolle wie dem wie dem Data Analyst oder Financial Analyst.

Mehr als **96.000** IT-Fachkräfte werden derzeit in Deutschland gesucht. Data Scientists bzw. Big Data Experts zählen dabei zu den meistgefragten Qualifikationen.



## Lerninhalte

---



Ziel des praxisorientierten Trainings ist es, Datenmodelle mit Python zur Vorhersage von unterschiedlichen Businesssszenarien umzusetzen. Du lernst, überwachte und unüberwachte Machine-Learning Algorithmen zu implementieren und zu optimieren. Ebenso baust Du Kompetenzen für gängige Visualisierungsmethoden auf, um Dich mit Abschluss des Data Scientist Trainings für die Jobrolle als Data Scientist zu qualifizieren.

Im Überblick:

- Daten selbständig einlesen, bereinigen und filtern
- Daten explorativ mithilfe von deskriptiver Statistik analysieren
- Komplexe Prognosemodelle entwickeln und verifizieren
- Python-Programmierkenntnisse vertiefen
- Datenmodelle zur Vorhersage von Businesssszenarien bauen
- Machine-Learning-Algorithmen entwickeln
- Best Practices für die effektive Datenvisualisierung
- Methoden des Data Storytellings

### ZIELGRUPPE:

Die Weiterbildung Data Scientist - Fokus Python eignet sich für alle, die Python als Programmiersprache einsetzen, Daten analysieren und auf Grundlage dieser Vorhersagen erstellen möchten, um datengetriebene Entscheidungen zu treffen. Du solltest eine grundlegende Motivation für Statistik, logisches Denken und maschinelles Lernen mitbringen. Auch für Quereinsteiger:innen ist die Data Scientist Weiterbildung geeignet.

## Kursüberblick

---



### DAUER:

108 Stunden (4,5 Monate)

### TYP:

Online-Training

### AUFBAU:

4 Module + 1 Abschlussprojekt

### NIVEAU:

Fortgeschrittene

### SPRACHE:

Deutsch, Englisch

### ABSCHLUSS:

Abschlusszertifikat

### VORAUSSETZUNGEN:

Für das Training werden Programmierkenntnisse in Python sowie Vorkenntnisse in den Python-Bibliotheken Pandas und Matplotlib vorausgesetzt.



# Modulübersicht

---

## MODUL 0: PREPARATION

### Ziel:

Auffrischung der Kenntnisse im Umgang mit Python sowie mathematischer Grundlagen Beschreibung

### Beschreibung:

Du führst Analysen und Datenmanipulationen in Python aus und nutzt dabei die Pakete Pandas und Matplotlib.

### Kapitel 1 – Data Analytics with Python:

Du machst Dich mit unserer interaktiven Programmierumgebung – dem Data Lab – vertraut und frischst die wichtigsten Programmier- und Python-Grundlagen zur Datenverarbeitung mit Pandas, Datenvisualisierung mit Matplotlib und Seaborn und Datenbankabfrage mit SQL Alchemy auf.

### Kapitel 2 – Linear Algebra:

Du machst Dich mit dem mathematischen Hintergrund von Data-Science-Algorithmen vertraut und lernst die Grundbegriffe der linearen Algebra kennen. Unter Verwendung des Pakets Numpy rechnest du mit Vektoren und Matrizen.

### Kapitel 3 – Probability Distributions:

Du lernst mehr über den statistischen Hintergrund von Data-Science-Algorithmen. Du beschäftigst Dich mit wichtigen statistischen Konzepten und lernst diskrete und kontinuierliche Verteilungen kennen.

## MODUL 1: MACHINE LEARNING BASICS

### Ziel:

Lösen von überwachten und unüberwachten Machine-Learning-Problemen mit sklearn

### Beschreibung:

Du erstellst Data-Science-Workflows mit sklearn, evaluierst ihre Modell-Performance anhand von geeigneten Metriken und wirst für das Problem des Overfittings sensibilisiert.

### Kapitel 1 – Supervised Learning (Regression):

Anhand der linearen Regression erlernst Du den Umgang mit dem Python-Paket sklearn. Weiterhin beschäftigst Du Dich mit den Annahmen des Regressionsmodells und der Evaluation der erzeugten Prognosen. In diesem Zuge werden auch der Bias-Variance Trade-Off, Konzepte der Regularisierung sowie verschiedene Maße der Modellgüte verdeutlicht.

### Kapitel 2 – Supervised Learning (Classification):

Du wirst in Klassifizierungsalgorithmen anhand des k-Nearest-Neighbors-Algorithmus eingeführt und lernst, den Algorithmus zu evaluieren und die Klassifizierungsperformance einzuschätzen. Du optimierst die Parameter ihres Modells unter Beachtung der Aufteilung der Daten in Trainings- und Evaluationssets.

### Kapitel 3 – Unsupervised Learning (Clustering):

Du lernst den k-Means-Algorithmus als Beispiel eines Algorithmus des unüberwachten Lernens kennen. Die Annahmen und Performance-Metriken des Algorithmus werden kritisch beleuchtet und ein kurzer Ausblick auf eine Alternative zum k-Means-Clustering geworfen.

**Kapitel 4 – Unsupervised Learning (Dimensionality Reduction):**

Du lernst, wie du mithilfe einer Principal Component Analysis (PCA) die Dimension der Daten verringern kannst und nutzt die PCA, um unkorrelierte Features aus den Ursprungsdaten zu erzeugen. In diesem Zusammenhang wird das Thema Feature Engineering näher betrachtet und aus den alten Features neue erzeugt.

**Kapitel 5 – Outlier Detection:**

Du lernst verschiedene Ansätze kennen, um Ausreißer zu identifizieren und verstehst, mit diesen ungewöhnlichen Datenpunkten umzugehen. Du nutzt robuste Maße und Modelle, um den Einfluss der Ausreißer zu minimieren.

**MODUL 2: DEEP DIVE SUPERVISED LEARNING****Ziel:**

Erweiterung des eigenen Data-Science-Toolkits

**Beschreibung:**

Du intensivierst Deine Kenntnisse über Modelle zur Klassifikation von Daten. Dabei erweiterst Du Deine Fähigkeiten im Sammeln und Aufbereiten von Daten.

**Kapitel 1 – Data Gathering:**

Du lernst, Daten zu sammeln, indem du Webseiten und PDF-Dokumente ausliest. Mithilfe von Regular Expressions strukturierst du gesammelte Textdaten so, dass du diese zusammen mit bekannten Algorithmen verwenden kannst.

**Kapitel 2 – Logistic Regression:**

Du lernst einen zweiten Klassifizierungsalgorithmus kennen: die logistische Regression. Du nutzt neue Performance-Metriken zur Evaluation der Ergebnisse und erfährst, wie du nicht-numerische Daten für Deine Modelle nutzbar machst.

**Kapitel 3 – Decision Trees and Random Forests:**

Du lernst den Entscheidungsbaum als leicht zu interpretierendes Modell kennen. Du kombinierst mehrere Modelle zu einem Ensemble, um die Vorhersagen Deines Modells zu verbessern. Weiterhin erhältst Du Methoden zu unausgeglichenen Kategorien an die Hand.

**Kapitel 4 – Support Vector Machines:**

Du lernst einen letzten Klassifizierungsalgorithmus kennen – Support Vector Machines (SVM) und beleuchtest das Verhalten verschiedener Kernel für die SVM. Außerdem erlernst Du die typischen Schritte des Natural Language Processing (NLP) und bearbeitest ein NLP-Szenario unter Verwendung von Bag-of-Words-Modellen.

**Kapitel 5 – Neural Networks:**

Du wirst in künstliche neuronale Netze eingeführt und lernst mehr über Deep Learning, um ein künstliches neuronales Netzwerk mit mehreren Schichten zu erzeugen und auf ein Datenset anzuwenden.

**MODUL 3: ADVANCED TOPICS IN DATA SCIENCE****Ziel:**

Selbstständiges Anwenden einfacher und komplexer Modellierungen

**Beschreibung:**

Du erlangst Souveränität im Lösen von Data-Science-Problemen und lernst, Ergebnisse kompetent zu kommunizieren.

**Kapitel 1 – Visualization and Model Interpretation:**

Du erlernst wichtige Methoden zur Interpretation und Visualisierung von Machine-Learning-Modellen. Durch die Verwendung modelagnostischer Methoden zur Interpretation lernst Du Erkenntnisse zur Funktionsweise ihrer Modelle abzuleiten und zu kommunizieren.

**Kapitel 2 – Spark:**

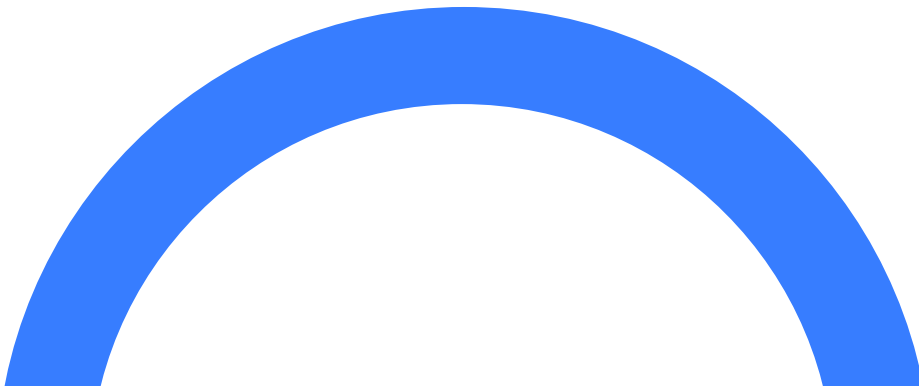
Du erfährst, weshalb die Arbeit mit verteilten Speichersystemen relevant ist. Mit dem Python-Paket PySpark erlernst Du verteilte Datenbanken auszulesen, Big-Data-Analysen durchzuführen und bekannte Machine-Learning-Algorithmen auf verteilten Systemen zu nutzen.

**Kapitel 3 – Exercise Project:**

Du bearbeitest ein Prädiktionsproblem mit Hilfe eines größeren Datensets und setzt Deine Data-Science-Fähigkeiten von der Reinigung des Datensets bis zur Interpretation des Modells eigenständig ein. In einer Projektbesprechung mit dem Mentoring Team von StackFuel erhältst Du Feedback zu Deinem Lösungsansatz.

**Kapitel 4 – Final Project:**

Du erhältst ein weiteres größeres Datenset, das Du selbstständig analysierst und im Vergleich zum Übungsprojekt mit weniger Hilfestellungen lösen musst. In einer individuellen Projektbesprechung mit dem Mentoring Team von StackFuel erhältst Du Feedback zu Deinem Lösungsansatz.

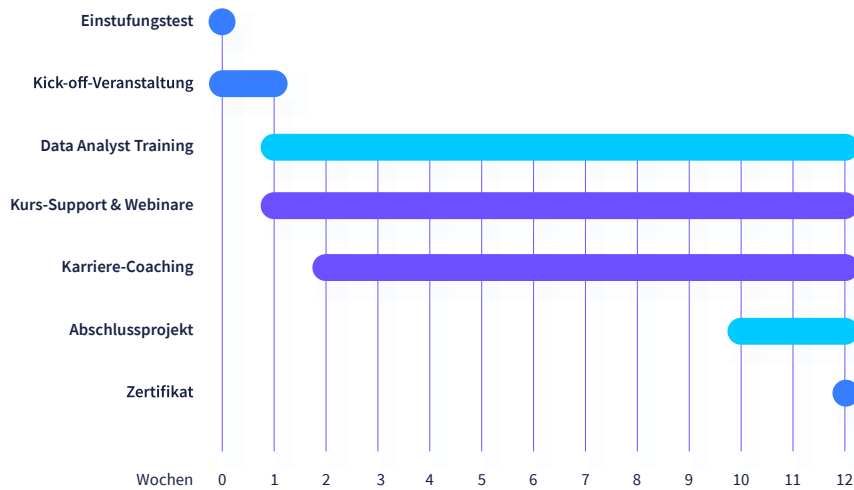




# Über StackFuel

Wir sind Dein strategischer Lernpartner, passend für jede Karrierestufe und berufliche Ausrichtung. Ob Quereinsteiger:innen, Mitarbeitende oder Führungskräfte mit unseren zertifizierten und staatlich geförderten Weiterbildungen im Bereich Daten, KI und Programmieren bleibst Du immer auf dem neusten Stand mit den aktuellen Technologien.

# Trainingsablauf



\*Exemplarischer Trainingsablauf für das Data Analyst Training.

- Vor- & Nachbereitung
- Mentoring & Support
- Trainingsinhalte

# Trainingsphilosophie



## ONLINE & FLEXIBEL

Mache Deine Weiterbildung in Teilzeit oder Vollzeit und lerne 100 % online im Browser über den Laptop oder PC.



## PRAXISNAH MIT EIGENEN PROJEKTEN

Im Data Lab schreibst Du in Praxisprojekten und Coding Challenges selbst Algorithmen mit Industriedatensets.



## MENTORING & KARRIERE-COACHING

Dein persönliches Mentoring Team begleitet Dich mit Beratung, Karriere-Coaching und in wöchentlichen Webinaren.



## ZERTIFIZIERT & FÖRDERFÄHIG

Als zertifizierter Träger kannst Du unsere staatlich geprüften Kurse über diverse Förderprogramme finanzieren lassen.

**80 %** interaktive Praxisinhalte  
**20 %** Expertenvideos & Textlektionen

**+150.000** Lernstunden im Data Lab  
**91 %** Abschlussquote  
**4,3 von 5** Weiterempfehlungsrate



# Fördermöglichkeiten

## RATEN- ODER TEILZAHLUNG

Nutze unsere Raten- oder Teilzahlungsoptionen, um die Kosten für Deine Weiterbildung auf mehrere Monate zu verteilen und Deine finanzielle Flexibilität zu erhalten.

## BILDUNGSGUTSCHEIN

Mit einem [Bildungsgutschein](#) kannst Du Dir Deine Weiterbildung zu 100 % durch das Jobcenter oder die Agentur für Arbeit finanzieren lassen, wenn Du aktuell arbeitslos oder arbeitssuchend registriert bist.

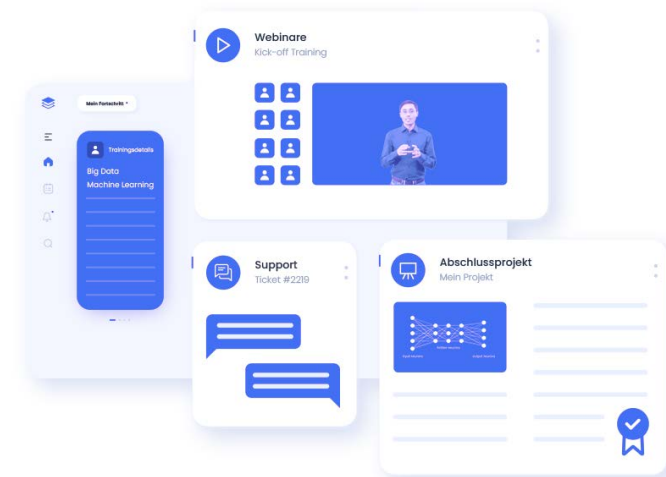
## QUALIFIZIERUNGSCHANCENGESETZ

Wenn Du berufstätig bist, kannst Du Deine Weiterbildung vom Arbeitgeber teilweise oder ganz über das [Qualifizierungschancengesetz](#) finanzieren lassen – unabhängig von Qualifikation, Alter und Betriebsgröße.

## STACKFUEL STIPENDIEN

Wir vergeben regelmäßig diverse [Stipendien](#) für unsere Trainings, um für mehr Diversität in der Datenbranche zu sorgen und immer mehr Menschen für das Programmieren zu begeistern und Frauen für datengetriebene Jobs zu begeistern.

# Support & Mentoring



## COACHING

- ✓ Einstufungstest
- ✓ Abschlussprojekt und -evaluation
- ✓ 1:1 Projekt-Feedbackgespräche
- ✓ Offizielles Abschlusszertifikat

## PERSÖNLICHE BETREUUNG

- ✓ Kick-off-Veranstaltung
- ✓ Webinare mit anderen Kursteilnehmenden
- ✓ Support via E-Mail oder Telefon
- ✓ Online-Forum mit Lernenden

## KARRIERE-SERVICES

- ✓ Lebenslauf- und Bewerbungstraining
- ✓ Talent-Pool und Karriere-Intros
- ✓ Data Community
- ✓ Karriere-Events





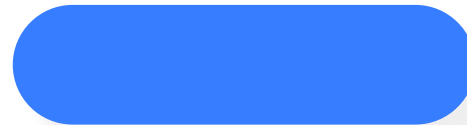
# Neugierig geworden?

**Melde Dich bei unseren Berater:innen und baue die Skills auf, die Du brauchst!**

## FÜR PRIVATPERSONEN



**Maria Schwenke**  
Weiterbildungsberaterin



## FÜR UNTERNEHMEN



**Ginesh Koottakara**  
Head of Sales

