




Kursstart alle 4 Wochen


# Big Data Analyst

Der Lehrgang zeigt die Anforderungen von Daten und Datenbanken, die Nutzung von Künstlicher Intelligenz (KI) in diesem Bereich, die Data Warehouse Modellierung, den ETL-Prozess und das Management der Datenanalyse im Big Data. Abschließend werden die Anforderungen des Data Engineerings erläutert.

 **Abschlussart**  
Zertifikat „Big Data Analyst“

 **Abschlussprüfung**  
Praxisbezogene Projektarbeiten mit Abschlusspräsentationen

 **Dauer**  
12 Wochen

 **Unterrichtszeiten**  
Montag bis Freitag von 8:30 bis 15:35 Uhr  
(in Wochen mit Feiertagen von 8:30 bis 17:10 Uhr)

 **Nächste Kursstarts**  
14.10.2024  
11.11.2024  
09.12.2024

## LEHRGANGSZIEL

Du beherrscht die Prozesse rund um die Zusammenführung, Aufbereitung, Anreicherung und Weitergabe von Daten und verstehst Big Data Analysen mit Hilfe grundlegender Python-Programmierungen, SQL und NoSQL-Datenbankkonzepten. Kenntnisse in branchenspezifischer Software zur Verarbeitung und Strukturierung von großen, unstrukturierten Daten sowie deren Visualisierung runden dein Wissen ab.

## ZIELGRUPPE

Der Lehrgang richtet sich an Personen mit abgeschlossenem Studium in der Informatik, Wirtschaftsinformatik, BWL, Mathematik oder vergleichbarer Qualifikation.

## BERUFSAUSSICHTEN

Eine systematische Auswertung von Datenmengen ist für Unternehmen unabdingbar, um Informationen über das eigene Produkt und Kundenverhalten zu generieren. Big Data Analysts werden vor diesem Hintergrund branchenübergreifend zunehmend nachgefragt.

Dein aussagekräftiges Zertifikat gibt detaillierten Einblick in deine erworbenen Qualifikationen und verbessert deine beruflichen Chancen.

## VORAUSSETZUNGEN

Programmierkenntnisse (idealerweise Python) und Erfahrungen mit Datenbanken (SQL) werden vorausgesetzt.

## LEHRGANGSINHALTE

### DATA ENGINEER

#### Grundlagen Business Intelligence (ca. 2 Tage)

Anwendungsfelder, Dimensionen einer BI Architektur  
Grundlagen Business Intelligence, OLAP, OLTP, Aufgaben der Data Engineers  
Data Warehousing (DWH): Umgang und Verarbeitung von strukturierten, semi-strukturierten und unstrukturierten Daten

#### Anforderungsmanagement (ca. 2 Tage)

Aufgaben, Ziele und Vorgehensweise in der Anforderungsanalyse  
Datenmodellierung, Einführung/Modellierung mit ERM  
Einführung/Modellierung in der UML  
· Klassendiagramme  
· Use-Case Analyse  
· Aktivitätsdiagramme

#### Künstliche Intelligenz (KI) im Arbeitsprozess

Vorstellung von konkreten KI-Technologien im beruflichen Umfeld  
Anwendungsmöglichkeiten und Praxis-Übungen

#### Datenbanken (ca. 3 Tage)

Grundlagen von Datenbanksystemen  
Architektur von Datenbankmanagementsystemen  
Anwendung RDBMS  
Umsetzung Datenmodell in RDBMS, Normalformen  
Praktische und theoretische Einführung in SQL  
Grenzen von Relationalen Datenbanken, csv, json

### Data Warehouse (ca. 4 Tage)

Star Schema  
Datenmodellierung  
Erstellung Star Schema in RDBMS  
Snowflake Schema, Grundlagen, Datenmodellierung  
Erstellung Snowflake Schema in RDBMS  
Galaxy Schema: Grundlagen, Datenmodellierung  
Slowly Changing Dimension Tables Typ 1 bis 5 – Restating, Stacking, Reorganizing, mini Dimension und Typ 5  
Einführung in normal, causal, mini und monster, heterogeneous und sub Dimensions  
Vergleich von state und transaction oriented Faktentabellen, Density und Storage vom DWH

### ETL (ca. 4 Tage)

Data Cleansing  
· Null Values  
· Aufbereitung von Daten  
· Harmonisierung von Daten  
· Anwendung von Regular Expressions  
Data Understanding  
· Datenvalidierung  
· Statistische Datenanalyse  
Datenschutz, Datensicherheit  
Praktischer Aufbau von ETL-Strecken  
Data Vault 2.0, Grundlagen, Hubs, Links, Satellites, Hash Key, Hash Diff.  
Data Vault Datenmodellierung  
Praktischer Aufbau eines Data Vault Modells – Raw Vault, Praktische Umsetzung von Hash-Verfahren

### Projektarbeit (ca. 5 Tage)

Zur Vertiefung der gelernten Inhalte  
Präsentation der Projektergebnisse

---

## DATA ANALYTICS

### Einführung Datenanalyse (ca. 1 Tag)

CRISP-DM Referenzmodell  
Data Analytics Workflows  
Begriffsabgrenzung Künstliche Intelligenz, Machine Learning, Deep Learning  
Anforderungen und Rolle im Unternehmen der Data Engineers, Data Scientists und Data Analysts

### Wiederholung Grundlagen Python (ca. 1 Tag)

Datentypen  
Funktionen

### Datenanalyse (ca. 3 Tage)

Zentrale Python-Module im Kontext Data Analytics (NumPy, Pandas)  
Prozess der Datenaufbereitung  
Data Mining Algorithmen in Python

### Künstliche Intelligenz (KI) im Arbeitsprozess

Vorstellung von konkreten KI-Technologien im beruflichen Umfeld  
Anwendungsmöglichkeiten und Praxis-Übungen

### Datenvisualisierung (ca. 3 Tage)

Explorative Datenanalyse  
Insights  
Datenqualität  
Nutzenanalyse  
Visualisierung mit Python: Matplotlib, Seaborn, Plotly Express  
Data Storytelling

### Datenmanagement (ca. 2 Tage)

Big Data Architekturen  
Relationale Datenbanken mit SQL  
Vergleich von SQL- und NoSQL-Datenbanken  
Business Intelligence  
Datenschutz im Kontext der Datenanalyse

### Datenanalyse im Big Data Kontext (ca. 1 Tag)

MapReduce-Ansatz  
Spark  
NoSQL

### Dashboards (ca. 3 Tage)

Bibliothek: Dash  
Aufbau von Dashboards – Dash Components  
Customizing von Dashboards  
Callbacks

### Text Mining (ca. 1 Tag)

Data Preprocessing  
Visualisierung  
Bibliothek: SpaCy

### Projektarbeit (ca. 5 Tage)

Zur Vertiefung der gelernten Inhalte  
Präsentation der Projektergebnisse

---

## BIG DATA SPECIALIST

### Was ist Big Data? (ca. 1 Tag)

Volume, Velocity, Variety, Value, Veracity  
Chancen und Risiken großer Datenmengen  
Abgrenzung: Business Intelligence, Data Analytics, Data Science  
Was ist Data Mining?

### Einführung in Apache Frameworks (ca. 2 Tage)

Big-Data-Lösungen in der Cloud  
Datenzugriffsmuster  
Datenspeicherung

### MapReduce (ca. 3 Tage)

MapReduce Philosophie  
Hadoop Cluster  
Verketten von MapReduce Jobs

### Künstliche Intelligenz (KI) im Arbeitsprozess

Vorstellung von konkreten KI-Technologien im beruflichen Umfeld  
Anwendungsmöglichkeiten und Praxis-Übungen

### Komponenten (ca. 3 Tage)

Kurzvorstellung von verschiedenen Tools  
Datenübertragung  
YARN-Anwendungen  
Hadoop JAVA-API  
Apache Spark

### NoSQL und HBase (ca. 3 Tage)

CAP-Theorem  
ACID und BASE  
Typen von Datenbanken  
HBase

### Big Data Visualisierung (ca. 3 Tage)

Theorien der Visualisierung  
Diagrammauswahl  
Neue Diagrammart  
Werkzeuge zur Datenvisualisierung

### Projektarbeit (ca. 5 Tage)

Zur Vertiefung der gelernten Inhalte  
Präsentation der Projektergebnisse

## UNTERRICHTSKONZEPT

### Didaktisches Konzept

Deine Dozierenden sind sowohl fachlich als auch didaktisch hoch qualifiziert und werden dich vom ersten bis zum letzten Tag unterrichten (kein Selbstlernsystem).

Du lernst in effektiven Kleingruppen. Die Kurse bestehen in der Regel aus 6 bis 25 Teilnehmenden. Der allgemeine Unterricht wird in allen Kursmodulen durch zahlreiche praxisbezogene Übungen ergänzt. Die Übungsphase ist ein wichtiger Bestandteil des Unterrichts, denn in dieser Zeit verarbeitest du das neu Erlernte und erlangst Sicherheit und Routine in der Anwendung. Im letzten Abschnitt des Lehrgangs findet eine Projektarbeit, eine Fallstudie oder eine Abschlussprüfung statt.

### Virtueller Klassenraum alfaview®

Der Unterricht findet über die moderne Videotechnik alfaview® statt - entweder bequem von zu Hause oder bei uns im Bildungszentrum. Über alfaview® kann sich der gesamte Kurs face-to-face sehen, in lippensynchroner Sprachqualität miteinander kommunizieren und an gemeinsamen Projekten arbeiten. Du kannst selbstverständlich auch deine zugeschalteten Trainer:innen jederzeit live sehen, mit diesen sprechen und du wirst während der gesamten Kursdauer von deinen Dozierenden in

Echtzeit unterrichtet. Der Unterricht ist kein E-Learning, sondern echter Live-Präsenzunterricht über Videotechnik.

## FÖRDERMÖGLICHKEITEN

Alle Lehrgänge werden von der Agentur für Arbeit gefördert und sind nach der Zulassungsverordnung AZAV zertifiziert. Bei der Einreichung eines Bildungsgutscheines oder eines Aktivierungs- und Vermittlungsgutscheines werden in der Regel die gesamten Lehrgangskosten von Ihrer Förderstelle übernommen.

Eine Förderung ist auch über den Europäischen Sozialfonds (ESF), die Deutsche Rentenversicherung (DRV) oder über regionale Förderprogramme möglich. Als Zeitsoldat:in besteht die Möglichkeit, Weiterbildungen über den Berufsförderungsdienst (BFD) zu besuchen. Auch Firmen können ihre Mitarbeiter:innen über eine Förderung der Agentur für Arbeit (Qualifizierungschancengesetz) qualifizieren lassen.

- ① Änderungen möglich. Die Lehrgangsinhalte werden regelmäßig aktualisiert. Die aktuellen Lehrgangsinhalte findest Du immer unter [smartbuilding.alfatraining.de](https://smartbuilding.alfatraining.de).