




Kursstart alle 4 Wochen

## Data Analyst


Data Analysts verfügen über Kenntnisse in der Statistik, in der Nutzung von Datenbanken und beherrschen die Programmiersprache Python. Auch verfügen sie über Fachwissen des Data Engineerings und der Datenanalyse und können dieses mit dem Einsatz von Künstlicher Intelligenz (KI) verbinden.

 **Abschlussart**  
Zertifikat „Data Analyst“

 **Abschlussprüfung**  
Praxisbezogene Projektarbeiten mit Abschlusspräsentationen

 **Dauer**  
20 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

Nach diesem Lehrgang verfügst du über wesentliche Kenntnisse in der Statistik, kannst komplexe Abfragen über relationale Datenbanken mit SQL durchführen und beherrscht die Programmiersprache Python. Verbunden mit dem im Kurs vermittelten Fachwissen des Data Engineerings und der Datenanalyse, bist du in der Lage, umfangreiche Datensätze zu managen, statistisch effizient auszuwerten und die Ergebnisse anschaulich und leicht verständlich zusammenzufassen.

### ZIELGRUPPE

Der Lehrgang richtet sich an Personen mit abgeschlossenem Studium in der BWL, Mathematik oder (Wirtschafts-)Informatik und vergleichbarer Qualifikation.

### BERUFSAUSSICHTEN

Da Unternehmen für die Auswertung und Zielsetzung Ihrer Geschäftsprozesse immer größer werdende Datenmengen bewältigen und strukturieren müssen, sind Kenntnisse im Bereich der Datenanalyse in allen Branchen nachgefragt.

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

### LEHRGANGSINHALTE

#### STATISTIK

##### Statistische Grundlagen (ca. 6 Tage)

Messtheoretische Grundlagen (Grundgesamtheit und Stichprobe, Stichprobenarten, Messung und Skalenniveaus)  
Univariate Deskriptivstatistik (Häufigkeitsverteilungen, Zentralmaße, Streuungsmaße, Standardwert, Histogramme, Balkendiagramme, Kreisdiagramme, Liniendiagramme und Boxplots)  
Bivariate Deskriptivstatistik (Zusammenhangsmaße, Korrelationskoeffizienten, Kreuztabellen, Streudiagramme und gruppierte Balkendiagramme)  
Grundlagen der induktiven Inferenzstatistik (Wahrscheinlichkeitsverteilung, Normalverteilung, Mittelwertverteilung, Signifikanztest, Nullhypothesentest nach Fisher, Effektgröße, Parameterschätzung, Konfidenzintervalle, Fehlerbalkendiagramme, Poweranalysen und Ermittlung des optimalen Stichprobenumfangs)

##### Künstliche Intelligenz (KI) im Arbeitsprozess

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

##### Methoden zum Vergleich von zwei Gruppen (ca. 5 Tage)

z- und t-Test für eine Stichprobe (Abweichung von einem vorgegebenen Wert)  
t-Test für den Mittelwertsunterschied von zwei unabhängigen/verbundenen Stichproben  
Prüfung der Wirksamkeit von Aktionen, Maßnahmen, Interventionen und anderen Veränderungen mit t-Tests (Pretest-Posttest-Designs mit zwei Gruppen)  
Unterstützende Signifikanztests (Anderson-Darling-Test, Ryan-Joiner-Test, Levene-Test, Bonnet-Test, Signifikanztest für Korrelationen)  
Nonparametrische Verfahren (Wilcoxon-Test, Vorzeichentest, Mann-Whitney-Test)  
Kontingenzanalysen (Binomialtest, Exakter Test nach Fisher, Chi-Quadrat-Test, Kreuztabellen mit Assoziationsmaße)

### Methoden zum Mittelwertvergleich von mehreren Gruppen (ca. 5 Tage)

Ein- und zweifaktorielle Varianzanalyse (einfache und balancierte ANOVA)  
Mehrfaktorielle Varianzanalyse (Allgemeines lineares Modell)  
Feste, zufällige, gekreuzte und geschachtelte Faktoren  
Mehrfachvergleichsverfahren (Tukey-HSD, Dunnett, Hsu-MCB, Games-Howell)  
Interaktionsanalyse (Analyse von Wechselwirkungseffekten)  
Trennschärfe und Poweranalyse bei Varianzanalysen

### Einführung in die Versuchsplanung (DoE, Design of Experiments) (ca. 1 Tag)

Voll- und teilfaktorielle Versuchspläne

### Projektarbeit (ca. 3 Tage)

Zur Vertiefung der gelernten Inhalte  
Präsentation der Projektergebnisse

---

## RELATIONALE DATENBANKEN MIT SQL

### Grundlagen von Datenbanksystemen mit Access (ca. 3 Tage)

Redundante Daten  
Datenintegrität  
Normalisierung  
BCNF  
DB-Entwurf  
Beziehung 1:n, m:n  
Datentypen  
Tabellen  
Primär- und Fremdschlüssel  
Referentielle Integrität  
Beziehungen zwischen Relationen  
Entity-Relationship-Modell  
Index, Standardwert  
Einschränkungen (Check)  
Abfragen  
Formulare, Berichte  
Zirkelbezug

### Einführung in SQL Server Management Studio (SSMS) (ca. 2 Tage)

Übersicht  
Phys. DB-Design  
Tabellen erstellen  
Datentypen in MS SQL  
Primary Key  
Einschränkungen, Standardwerte, Diagramm, Beziehungen  
Backup und Restore

### Künstliche Intelligenz (KI) im Arbeitsprozess

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

### Einführung in DDL (ca. 8 Tage)

SQL Grundlagen  
Syntax  
Befehle  
Mehrere Tabellen  
Operatoren  
Ablaufkontrolle  
Skalarwertfunktionen  
Tabellenwertfunktionen  
Systemfunktionen  
Prozeduren mit und ohne Parameter  
Fehlertypen  
Transaktionen, Sperren, DeadLock

### DCL – Data Control Language (ca. 1 Tag)

Anmeldungen  
Benutzer:innen  
Rollen  
Berechtigungen

### Datentypen, Datenimport und -export (ca. 1 Tag)

Datentyp geography  
Datenexport, Datenimport

### Projektarbeit (ca. 5 Tage)

Zur Vertiefung der gelernten Inhalte  
Präsentation der Projektergebnisse

---

## PROGRAMMIERUNG MIT PYTHON

### Grundlagen Python (ca. 1 Tag)

Geschichte, Konzepte  
Verwendung und Einsatzgebiete  
Syntax

### Erste Schritte mit Python (ca. 5 Tage)

Zahlen  
Zeichenketten  
Datum und Zeit  
Standardeingabe und -ausgabe  
list, tuple dict, set  
Verzweigungen und Schleifen (if, for, while)

### Künstliche Intelligenz (KI) im Arbeitsprozess

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

### Funktionen (ca. 5 Tage)

Eigene Funktionen definieren  
Variablen  
Parameter, Rekursion  
Funktionale Programmierung

### Fehlerbehebung (ca. 0,5 Tage)

try, except  
Programmunterbrechungen abfangen

### Objektorientierte Programmierung (ca. 4,5 Tage)

Python-Klassen  
Methoden  
Unveränderliche Objekte  
Datenklasse  
Vererbung

### Grafische Benutzeroberfläche (ca. 1 Tag)

Buttons und Textfelder  
grid-Layout  
Dateiauswahl

### Projektarbeit (ca. 3 Tage)

Zur Vertiefung der gelernten Inhalte  
Präsentation der Projektergebnisse

---

## 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

## 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).