

# Data Science mit Python - Einführung in Python (S1961)

Python ist eine der weitverbreitetsten Programmiersprachen, welche im Data Science Bereich etabliert ist. Neben R ist sie derzeit die meist benutzte Data Science Programmiersprache. Sie erhalten einen Einstieg, um mit Python Daten zu analysieren und Machine Learning Algorithmen umsetzen zu können.

## Offene Termine

Termin	Tage	Freie Plätze	Ort	Preis
25.11.-27.11.2019 ✓ <span style="border: 1px solid black; padding: 0 2px;">3=2</span>	3	>3	Köln	€ 1.730,00 *
13.01.-15.01.2020 ✓ <span style="border: 1px solid black; padding: 0 2px;">3=2</span>	3	>3	Köln	€ 1.730,00 *
15.04.-17.04.2020 ✓ <span style="border: 1px solid black; padding: 0 2px;">3=2</span>	3	>3	Köln	€ 1.730,00 *
15.06.-17.06.2020 ✓ <span style="border: 1px solid black; padding: 0 2px;">3=2</span>	3	>3	Köln	€ 1.730,00 *
17.08.-19.08.2020 ✓ <span style="border: 1px solid black; padding: 0 2px;">3=2</span>	3	>3	Köln	€ 1.730,00 *
14.09.-16.09.2020 ✓ <span style="border: 1px solid black; padding: 0 2px;">3=2</span>	3	>3	Köln	€ 1.730,00 *
09.11.-11.11.2020 ✓ <span style="border: 1px solid black; padding: 0 2px;">3=2</span>	3	>3	Köln	€ 1.730,00 *

### \* Buchen ohne Risiko

- › Keine Vorkasse
- › Kostenloses Storno bis zum Vortag des Seminars
- › Rechnung nach erfolgreichem Seminar

### ✓ Garantierter Termin und Veranstaltungsort

€ Preise zzgl. Mehrwertsteuer

3=2 Der dritte Mitarbeiter nimmt kostenlos teil

## Weitere Buchungsmöglichkeiten

<b>Firmenschulung</b>	Schulung für Ihre Mitarbeiter mit individuellen Inhalten zum Wunschtermin im GFU-Schulungszentrum.
<b>Inhouse-Schulung</b>	Schulung für Ihre Mitarbeiter mit individuellen Inhalten zum Wunschtermin in Ihrem Hause.
<b>Individualschulung</b>	Schulung für eine Einzelperson mit individuellen Inhalten zum Wunschtermin, wahlweise in Ihrem Hause oder im GFU-Schulungszentrum.

## Schulungs-Ziel

Sie können grundlegende Schritte mit Python im Bereich Data Science anwenden, so dass ein

## Inhalt

- › **Einführung**
  - › Kurzer Hintergrund zu Python und Anaconda

selbstständiges Arbeiten möglich ist. Wege, um weitergehende Algorithmen und Methoden in Python zu finden sind bekannt, um das Wissen auch nach dem Seminar weiter ausbauen zu können. Sie haben einen Überblick über verwendete Algorithmen im Machine Learning und können diese voneinander differenzieren und in Python anwenden. Sie arbeiten überwiegend selbstständig mit Unterstützung des Trainers, um das Erlernte direkt umzusetzen und anzuwenden. Anhand eines durchgehenden Datenbeispiels werden die wichtigsten Bibliotheken besprochen und in umfangreichen Übungen trainiert.

## Wer sollte teilnehmen

---

Technisch interessierte Fachkräfte bzw. Projektleiter, welche einen Einstieg in die Programmiersprache Python wünschen, um eigenständig an Data Science Projekten mitzuarbeiten oder ein besseres Verständnis für Python Code zu erhalten. Die Teilnehmer benötigen keine Erfahrung mit Python, sollten aber schon grundlegende Erfahrung mit einer Programmiersprache gemacht haben, um das Grundkonzept einer Programmiersprache zu verstehen. Überdies ist ein Basiswissen von grundlegenden statistischen Begriffen hilfreich (wie Mittelwert, Median, Perzentil).

## Organisation

---

### Teilnehmerzahl

min. 1, max. 8 Personen

### Seminarzeiten

3 Tage, 1. Tag 10:00 - 17:00 Uhr, Folgetage 09:00 - 16:00 Uhr

### Ort der Schulung

GFU-Schulungszentrum Köln oder bei Ihnen als Inhouse-Schulung

## Enthaltene Leistungen

---

### Im Preis enthalten:

- > Voll ausgestatteter Arbeitsplatz pro Teilnehmer
- > Fachbuch zum Seminar
- > Teilnahmezertifikat

- > Installieren von neuen Paketen
- > Die Entwicklungsumgebung (IDE) Spyder
- > Aspekte von Python, die anders als in anderen Programmiersprachen sind
- > **Datenstrukturen in Python**
  - > Unterschied Funktionen vs. Methoden
  - > Datenstrukturen in Python (number, tuple, list, dictionary, string)
  - > Überblick und wichtige Methoden zu den Datenstrukturen
  - > Unterschied list vs. tuple
  - > Die Besonderheit des Wertes NaN
  - > Wann benutze ich welche Klammer [], (), {}?
  - > List comprehension
- > **Einführung in das Data Science Paket pandas**
  - > Elemente eines pandas data.frame
  - > Eine Zeile ansprechen, hinzufügen, löschen und verändern
  - > Eine Spalte ansprechen, hinzufügen, löschen und verändern
  - > Boolean indexing
  - > Ein leeres data.frame erstellen
  - > Überblick über Methoden, ein data.frame anzupassen (Spaltennamen ändern, Werte ersetzen, ein data.frame sortieren, ein data.frame nach einer kategorischen Variable gruppieren, randomisiert Zeilen aus einem data.frame auswählen)
- > **Grundlegende Statistiken mit pandas**
  - > Überblick über Statistiken, die ich mit einem data.frame berechnen kann (Anzahl an gültigen Beobachtungen, Summe, Mittelwert, Median, Minimum, Maximum, Varianz,...)
  - > Eine Kreuztabelle (Kontingenztafel) berechnen
  - > Fehlende Werte ergänzen
  - > Logische Operatoren
- > **Daten einlesen**
  - > Das Arbeitsverzeichnis in Python setzen
  - > Ein CSV einlesen und schreiben und Überblick über die nützlichsten Parameter hierfür
  - > Python's pickle modul zum Speichern von Python Objekten
  - > Verschiedene Datentypen konvertieren (integer, unsigned integer, float, string)
- > **Visualisierung mit matplotlib und seaborn**
  - > Die Hauptelemente beim Plotten mit matplotlib
  - > Einen Plot anpassen (x- und y-Achse ändern, Beschriftungen, Legende und Titel ändern) und speichern
  - > Auswahl von Farben
  - > Subplots erstellen
  - > Erstellen von Scatterplot, Liniplot, Barplot, (gruppiertes) Histogram, Boxplot
- > **Control Flows**
  - > Die range() Funktion

- > Kostenloser persönlicher Parkplatz
- > Kostenloser Shuttle-Service
- > Frühstück, Snacks und Getränke ganztägig
- > Mittagessen im eigenen Restaurant, täglich 6 Menüs, auch vegetarisch

## Haben Sie Fragen?

---

Gerne beraten wir Sie persönlich per [Mail](#) oder Telefon.

- > [info@gfu.net](mailto:info@gfu.net)
- > Infoline 0221 82 80 90

- > Eine eigene Funktion schreiben
- > Default Parameter in einer Funktion setzen
- > For Schleifen
- > If-Else Bedingungen
- > While Schleife
- > Logische Vergleichsoperatoren
- > Zuweisungsoperatoren
- > **Überblick über Machine Learning**
  - > Was ist Machine Learning
  - > Die Einteilung von Algorithmen im Machine Learning (supervised Learning - unsupervised learning - reinforcement learning)
  - > Overfitting, Train-Test-Split und cross-validation (Kreuzvalidierung)
- > **Machine Learning Algorithmen (Theorie und Praxis)**
  - > Einführung in den Algorithmus (Lineare Regression, Entscheidungsbaum, Logistische)
  - > Train-Test Split der Daten
  - > Umsetzung des Algorithmus in Python
  - > Validieren der Ergebnisse (u.a. confusion matrix, sensitivity, accuracy)
  - > Cross-validation (Kreuzvalidierung)
- > **Weitere Machine Learning Algorithmen in Python**
  - > Erklärung von (SVM), Random Forest und K-means
  - > Umsetzung dieser Algorithmen in Python
  - > Validieren der Ergebnisse
  - > Cross-validation (Kreuzvalidierung)