Centrum Szkoleniowe Comarch
ul. Prof. Michała Życzkowskiego 33
31-864 Kraków

COMARCH
Szkolenia

Python w analizie danych

Wstęp do Data Science

Cele szkolenia

Celem szkolenia jest wprowadzenie uczestników w podstawowe zagadnienia dotyczące analizy danych z użyciem języka Python i najważniejszych bibliotek Pythona dedykowanych obszarowi Data Science (Pandas, Numpy, Matplotlib, Seaborn). Dzięki szkoleniu uczestnicy będą w stanie z łatwością wczytać, przeprocesować, wyczyścić i łączyć duże zbiory danych, a następnie przeprowadzić pełnowymiarową analizę wraz z prostymi wizualizacjami.

Umiejętności

Dzięki szkoleniu uczestnik będzie:

- stawiać odpowiednie pytania odnośnie danych i przeprowadzać w sposób świadomy proces analizy,
- posługiwać się przydatnymi w analizie danych bibliotekami języka Python,
- selekcjonować, oczyszczać i przetwarzać dane przy użyciu nowoczesnych technik,
- prezentować dane przy użyciu odpowiednio dobranych wykresów i wizualizacji.

Profil uczestników

Szkolenie jest przeznaczone dla osób chcących zapoznać się z możliwościami przetwarzania, analizy, eksploracji i wizualizacji danych z użyciem języka Python.

Szkolenie może być szczególnie interesujące dla osób zainteresowanych rozwojem na następujących stanowiskach: data analyst (analityk danych), business analyst (analityk biznesowy), reporting analyst (analityk ds. raportowania), data scientist, business intelligence BI analyst (analityk ds. business intelligence).

Przygotowanie uczestników

Od uczestników szkolenia wymagana jest znajomość podstaw programowania w języku Python (w tym tworzenie i uruchamianie kodu w środowisku Jupyter Notebook). Przydatne będzie doświadczenie w analizie danych w innym narzędziu np. Excel.

Szczegółowy program szkolenia

Wprowadzenie

- Omówienie celów i zakresu szkolenia
- Na czym polega analiza danych i czemu jest istotna

- Rola Pythona w analizie danych
- Omówienie środowiska pracy - efektywne korzystanie z Jupyter Notebook

Podstawy statystyki

- Podstawowe pojęcia statystyczne
- Podstawowe miary rozkładu

Python - wybrane elementy

- Przypomnienie podstawowych typów i struktur danych
- Indeksowanie, slicing
- Funkcje, funkcje anonimowe
- List comprehension
- Wybrane funkcje wbudowane: matematyczne, tekstowe, statystyczne, data i czas
- Kontrola przepływu (while, for, if, elif, else), przydatne rozszerzenia (enumerate, zip)

Analiza z NumPy i Pandas

- Tablice jedno i dwuwymiarowe w NumPy oraz podstawowe operacje
- Series i DataFrame w Pandas
- Wczytywanie i zapis danych w różnych formatach
- Podstawowe atrybuty DataFrame (shape, indeksy, kolumny, typy danych)
- Przydatne funkcje (describe, info, head, tail, sample)
- Czyszczenie wartości zduplikowanych
- Wartości brakujące - różne podejścia do radzenia sobie z nimi
- Wykrywanie wartości odstających
- Sortowanie danych
- Filtrowanie (loc, iloc, query, where, isin, isnull, notnull)
- Tabele przestawne
- Grupowanie
- Łączenie danych
- Tworzenie nowych atrybutów

Wizualizacja danych z Maplotlib i Seaborn

- Rodzaje wykresów i wizualizacji - dobór wykresu, elementy storytelling
- Podstawowe typy wykresów (histogram, wykres słupkowy, wykres typu scatter) w Pythonie

Metoda realizacji szkolenia

Szkolenie składa się z części teoretycznych oraz ćwiczeń praktycznych, a także case studies, czyli faktycznej analizy zbiorów danych z różnych obszarów. Szkolenie łączy w sobie wiedzę merytoryczną z praktycznymi przykładami jej wykorzystania w środowisku pracy analityka.

Liczba dni, liczba godzin szkoleniowych

2 dni, 16 godzin szkoleniowych

Ścieżka rozwoju po szkoleniu

Wstęp do Machine Learning i Deep Learning w języku Python
https://www.comarch.pl/szkolenia/programowanie/python/wstep-do-machine-learning-i-deep-learning/

Algorytmy Sztucznej Inteligencji (AI). Implementacja w języku Python
https://www.comarch.pl/szkolenia/programowanie/python/algorytmy-sztucznej-inteligencji-ai-implementacja-w-jezyku-python/

Machine Learning z użyciem języka Python. Zagadnienia zaawansowane
https://www.comarch.pl/szkolenia/programowanie/python/machine-learning-z-uzyciem-jezyka-python/