

Tytuł szkolenia: Analiza i wizualizacja danych w Pythonie

Kod szkolenia: analiza-i-wizualizacja-danych-python

Wprowadzenie

Python jest obecnie jednym z **najpopularniejszych języków programowania**. Język ten ma wiele zastosowań, lecz w ostatnich latach ogromną popularnością cieszy się jego wykorzystanie w obszarze analizy danych. Systematycznie **rośnie liczba firm wykorzystujących** Pythona oraz jego zastosowań w takich obszarach jak uczenie maszynowe, data science, sieci neuronowe, big data. W tym kontekście niezwykle ważne jest posiadanie solidnych podstaw analitycznych, bez których świadome i użyteczne wykorzystanie z Pythona w tych obszarach, będzie bardzo trudne. Już znajomość podstawowych zagadnień programistycznych i analitycznych pozwala wykonywać efektywne, prowadzące do wartościowych wniosków analizy.

Adresaci szkolenia

Zajęcia dedykowane są zarówno osobom **rozpoczynającym przygodę z programowaniem lub Pythonem**, jak również tym, którzy **znają jego podstawy**, ale chcieliby **usystematyzować wiedzę i umiejętności** niezbędne do przeprowadzania statystycznej i eksploracyjnej analizy danych w języku Python. Program kursu skonstruowany jest w taki sposób, aby każdy, niezależnie od tego, czy zna Pythona, czy nie, mógł zdobyć solidne fundamenty do samodzielnej pracy i dalszej edukacji obejmującej zagadnienia z zakresu uczenia maszynowego.

Cel szkolenia

Podstawowym celem kursu jest **zapoznanie uczestników z podstawami programowania w języku Python**, ze szczególnym uwzględnieniem zagadnień analitycznych. Tematyka kursu obejmuje podstawy pracy ze środowiskiem Python, podstawy teoretyczne oraz zagadnienia programistyczne, niezbędne do efektywnego przetwarzania danych. Równie ważnym elementem szkolenia będzie zapoznanie uczestników kursu z wybranymi zagadnieniami matematycznymi i statystycznymi, które pozwolą w sposób samodzielny i świadomy wykonywać prace analityczne i formułować wnioski płynące z analizy danych.

Kurs ma charakter **warsztatowy**, co oznacza, że główny nacisk położony będzie na omówienie praktycznych aspektów przeprowadzanych analiz. Po zakończeniu kursu, każdy uczestnik będzie posiadał solidne podstawy do prowadzenia w pełni samodzielnych analiz oraz zdobędzie niezbędną wiedzę i umiejętności, by przystąpić do bardziej zaawansowanych zagadnień analitycznych tj. uczenie maszynowe, Data Science.

Czas i forma szkolenia

- 21 godzin (3 dni x 7 godzin), w tym wykłady i warsztaty praktyczne.

Plan szkolenia

Przygotowanie środowiska do pracy

- Anaconda i Jupyter Notebook
- Środowisko wirtualne
- podstawy pracy z linią poleceń Windows/Anaconda

Podstawy programowania

- zmienne i ich typy
- operatory
- podstawowe struktury danych
- kontrola przepływu
- pętle
- funkcje

Wektorów i macierzy Numpy

- tworzenie obiektów
- wektory vs. listy
- przekształcenia
- indeksowanie
- łączenie i dzielenie macierzy
- sortowanie
- podstawy algebry liniowej
- przykłady funkcji statystycznych

Przetwarzanie danych z pakietem Pandas

- Data Series i Data Frame
- wczytywanie danych
- modyfikacja danych
- indeksowanie i filtrowanie
- operacje sortowania.
- łączenie obiektów
- podstawowe funkcje statystyczne
- grupowanie danych

Podstawy ilościowej i eksploracyjnej analizy danych z pakietem Pandas i Matplotlib

- statystyka opisowa
- korelacja zmiennych
- wykresy jednej zmiennej
- wykresy dwóch zmiennych
- case study