Analiza danych w Pythonie (bez ML) (kod: K-PYTHON-A-SHORT)

Zapytaj o szczegóły
tel. 22 63 64 164
akademia@alx.pl

Najbliższe terminy
2026-12-10 (Zdalnie)
2026-12-10 (Warszawa)
2026-12-12 (Zdalnie)
2026-12-12 (Warszawa)

Opis i cel kursu

Umiejętność odnajdowania i analizy zjawisk w dużych zbiorach danych to konieczność we współczesnym świecie. Sprawne przetwarzanie danych pozwala podejmować skuteczne i konkurencyjne decyzje - a to gwarantuje sukces w biznesie.

Nie każda firma potrafi wykorzystać potencjał jaki kryje się w danych. Aby analizować duże ilości, często niepełnych informacji konieczne są metody statystyczne i znajomość odpowiednich technologii. Te umiejętności zdobędziesz na naszym kursie!

— *Python** jest jednym z najpopularniejszych języków do **analizy danych** dzięki bogatemu zapleczu narzędzi.

— *pandas**, **NumPy**, **matplotlib** i inne popularne pakiety stanowią dojrzały ekosystem gotowych do użycia modułów, a uniwersalność Pythona pozwala poza samą analizą danych także je pobierać, obrabiać i eksportować zarówno w formie raportów, jak i plików wejściowych dla innych aplikacji.

Na naszym kursie osoby znające podstawy Pythona mogą poszerzyć wiedzę o zagadnienia związane z analizą danych, a także poznać narzędzia wykorzystywane w tym celu.

Od uczestników oczekiwana jest podstawowa znajomość Pythona.

Program

1. **Sposoby pracy z Pythonem okiem analityka danych**
   - — Instalacja lokalna
     - — Interpreter języka Python
     - — Tworzenie i uruchamianie programów
     - — Zintegrowane środowiska programistyczne (IDE)
   - — Konfiguracja lokalnej instalacji
     - — Wirtualne środowisko (`venv`) i instalator pakietów `pip`
     - — Instalacja Jupytera i dodatkowych bibliotek
   - — Anaconda jako alternatywny sposób instalacji lokalnej – podejście „wszystko w jednym”
     - — Zarządzanie wersjami bibliotek w Anacondzie
   - — Środowisko Jupyter
     - — Komórki i specyfika pracy interaktywnej
     - — Skróty klawiszowe i inne opcje ułatwiające pracę
     - — Podpowiedzi i dokumentacja
     - — Używanie Markdown i Latex do tworzenia sformatowanych tekstów
   - — Możliwości pracy zdalnej – Google Colab i podobne usługi
   - — Rzut oka na „ekosystem” bibliotek i narzędzi służących do *data science* w Pythonie
     - — Asystent AI Gemini
2. **Podstawy języka Python**
   - — Zmienne, wartości, typy
     - — Liczby i operacje arytmetyczne
     - — Napisy, w tym umieszczanie wartości w tekście za pomocą f-string
     - — Wartości logiczne
   - — Warunki, porównania, spójniki logiczne

- Definiowanie własnych funkcji

3. **Struktury danych języka Python**
   - Listy i krotki
     - Iteracja po elementach (pętla `for`)
     - Wybieranie elementów i zakresów (*slicing*)
   - Zbiory i słowniki
   - Wyrażenia generujące kolekcje (*comprehensions*)
   - Dostęp do danych zewnętrznych: pliki tekstowe, dane JSON
   - Funkcje matematyczne i statystyczne zawarte w bibliotece standardowej Pythona
4. **Numpy - tablice i obliczenia w Pythonie**
   - Przyczyny używania dodatkowych bibliotek analityczno-obliczeniowych
   - Wektory, macierze, tablice wielowymiarowe
     - Różne sposoby tworzenia
     - Nawigacja i wybieranie fragmentów
     - Zmiana kształtu (`reshape`) i transpozycja
   - Typy liczbowe i konsekwencje wyboru typu
   - Operacje na tablicach Numpy
     - Wektoryzacja
     - Zasada rozgłaszania (*broadcasting*)
     - Mnożenie macierzy i przykładowe zastosowania
     - Funkcje agregujące i osie
     - Rozkłady, histogramy, statystyki opisowe
   - Generator liczb pseudolosowych
5. **Pandas i „biznesowa" analiza danych**
   - Serie (`Series`) i tabele / „ramki danych" (`DataFrame`)
     - Indeksy i nazwy kolumn
     - Nawigacja w strukturach danych - wybieranie komórek i fragmentów
   - Pozyskiwanie danych z różnych źródeł
     - Pliki CSV i Excel (odczyt i zapis)
     - Zasoby w internecie (JSON, XML, HTML)
     - Bazy danych (SQL)
   - Wyszukiwanie i filtrowanie danych
     - Warunki logiczne w Pandas oraz Numpy: maski, spójniki logiczne
   - Przygotowywanie i czyszczenie danych
     - Usuwanie kolumn i wierszy
     - Usuwanie duplikatów
     - Zamiana i normalizacja wartości
     - Strategie postępowania z wartościami nieokreślonymi (NaN)
     - * Usuwanie wartości nieokreślonych
     - * Zastępowanie domyślną lub dominującą wartością
     - * Wypełnianie wartościami sąsiednimi
     - * Interpolacja
     - Praca z danymi tekstowymi, w tym wykorzystanie wyrażeń regularnych (*regex*)
   - Pionowe oraz poziome łączenie tabel (`concat`, `merge`, `join`)
   - Sortowanie serii i tabel
     - Ranking i poszukiwanie najmniejszych/największych wartości
   - Opisowa analiza danych
     - Gotowe funkcje do generowania podstawowych statystyk
     - Grupowanie, agregacja i obliczanie statystyk
     - Tabele przestawne (*pivot table*)
     - Funkcje kroczące („okienkowe") i narastające („skumulowane")
   - Oś czasu i szeregi czasowe

Zapytaj o szczegóły
tel. 22 63 64 164
akademia@alx.pl

Najbliższe terminy
2026-12-10 (Zdalnie)
2026-12-10 (Warszawa)
2026-12-12 (Zdalnie)
2026-12-12 (Warszawa)

- Generowanie szeregów czasowych
- Specyfikowanie okresów czasu
- Agregacja okresów czasu – operacja `resample`
- Elementy analizy statystycznej
  - Seria danych jako zmienna statystyczna
  - Korelacja, mapy ciepła
  - Czym jest regresja liniowa? (podstawy)

6. **Wykresy i wizualizacja danych**
   - Matplotlib i wykresy na podstawie danych z Pandas i Numpy
     - Różne typy wykresów
     - Opcje i ustawienia
   - Generowanie wykresów bezpośrednio z Pandas
   - Przegląd dodatkowych bibliotek wizualizacji danych, m.in. Seaborn, Bokeh, Plotly

Przeznaczenie i wymagania

Brak wymagań, kurs od podstaw. Jednak podstawowa znajomość Pythona będzie dodatkowym atutem.

Certyfikaty

Uczestnicy szkolenia otrzymują imienne certyfikaty sygnowane przez ALX.

Zapytaj o szczegóły

tel. 22 63 64 164
akademia@alx.pl

Najbliższe terminy

2026-12-10 (Zdalnie)
2026-12-10 (Warszawa)
2026-12-12 (Zdalnie)
2026-12-12 (Warszawa)