

kod szkolenia: PYTH_ANALIZA_ZAAW / PL AA 2d

Analiza danych w Pythonie – poziom średniozaawansowany

Szkolenie: Analiza danych w Pythonie – poziom
średniozaawansowany



Przeznaczenie szkolenia

Szkolenie przeznaczone jest dla osób chcących poznać techniki **analizy danych** z wykorzystaniem języka **Python** na poziomie średniozaawansowanym.



Korzyści wynikające z ukończenia szkolenia

Szkolenie stanowi kontynuację kursu „Analiza danych w Pythonie – poziom podstawowy”. Uczestnicy rozwiną dalsze umiejętności **wizualizacji danych** z użyciem **pakietu Seaborn** oraz zaznajomią się z narzędziem do interaktywnej wizualizacji danych – **pakiem Bokeh**. Poznają podstawy prawdopodobieństwa, statystyki oraz wnioskowania statystycznego. Kursanci nauczą się przeprowadzać symulacje losowe oraz testować hipotezy pozwalające określić, czy występujące w populacjach różnice są istotne statystycznie.



Oczekiwane przygotowanie słuchaczy

Znajomość podstaw języka Python oraz pakietów NumPy, pandas i Matplotlib.



Język szkolenia

- Szkolenie: polski



Czas trwania

2 dni / 14 godzin

Agenda szkolenia

1. Zaawansowana **wizualizacja danych**

- **Seaborn**
 - Heatmap
 - **Pair plot**
 - **Swarm plot**
 - Box plot
- Bokeh
 - **Interaktywna wizualizacja**
 - Konfiguracja wyglądu
 - Wykres liniowy
 - **Wykres rozrzutu**
 - Suwaki i przyciski

2. Podstawy prawdopodobieństwa i **statystyki**

- Prawdopodobieństwo
- **Korelacja danych**
- **Rozkład ciągły** i dyskretny
- **Gęstość prawdopodobieństwa**
- **Funkcja masy prawdopodobieństwa**
- Rozkład standardowy
- Centralne twierdzenie graniczne
- **Reguła 3 sigma**

3. **Wnioskowanie** statystyczne

- Liczby pseudolosowe
- Symulacje losowe
- **p-value**
- Estymacja punktowa
- Estymacja przedziałowa
- Bootstrapowy **przedział ufności**

- **Metoda Monte-Carlo**
- **Testowanie hipotez**
- **Test równości średnich**
- **Test równości** proporcji
- Test normalności rozkładu