Podstawy analizy danych w R (kod: R-INTRO)

Opis i cel szkolenia

To praktyczne i kompleksowe szkolenie przygotowaliśmy dla wszystkich kursantów, którzy chcą poznać tajniki analizy danych w szybko zyskującym popularność programie statystycznym R.

Wiedzę o programie R uzyskaną na początku szkolenia kursanci będą mogli od razu wykorzystać podczas nauki podstaw statystyki. Drugiego dnia skupimy się na regresji liniowej - jednej z podstawowych metod analizy stosowanej w małych i dużych firmach. Trzeciego dnia poszerzymy znajomość R i uczymy się wykorzystywać jego możliwości do prezentowania wyników w formie czytelnych wykresów.

Szkolenie prowadzone jest w formie warsztatów, gdzie teoria przeplata się z praktyką umożliwiając natychmiastowe przećwiczenie nowo poznanych narzędzi na życiowych, praktycznych przykładach.

Szkolenie jest dość intensywne i skupia się na warsztacie analityka - poznaniu środowiska R i wykorzystaniu go do przeprowadzenia przerabianych przykładów. Nie uczymy tutaj samych idei/fundamentów analizy danych od podstaw. Osobom, które dopiero zaczynają przygotowanie do pracy analityka, sugerujemy nasz dłuższy, pełny kurs Analiza danych w R.

Czas trwania

3 dni, 9:00 - 17:00

Program

1. **Wprowadzenie do R i RStudio**
   - R jako język programowania i RStudio jako wygodny interfejs użytkownika
   - Typy danych w R: skalar, wektor, macierz, ramka danych, factor, lista
   - R jako kalkulator: operatory arytmetyczne, porównawcze i logiczne, rodzaje poleceń: wyrażenia vs przypisania
   - Przetwarzanie danych: wybór elementów wektora/macierzy na podstawie indeksów, tworzenie nowych kolumn w ramce danych, wybór wierszy z ramki danych na podstawie warunków
   - Podstawowa analiza danych numerycznych (obliczanie średniej, sumy, itp.) i tekstowych (tabela częstości)
2. **Statystyka opisowa**
   - Rodzaje danych statystycznych: ilościowe vs jakościowe
   - Obliczanie i interpretacja statystyk opisowych dotyczących wartości przeciętnej, zróżnicowania oraz kształtu rozkładu zmiennej ilościowej
   - Nieparametryczne miary rozkładu: dominanta/moda, mediana, kwartyle, decyle, kwantyle
   - Podsumowanie rozkładu zmiennej jakościowej – tabele częstości
3. **Wnioskowanie statystyczne**
   - Popularne rozkłady prawdopodobieństwa, m.in. jednostajny, normalny, t-Studenta, Chi-kwadrat, F-Snedecora
   - Generowanie liczb (pseudo)losowych, określanie ziarna generatora
   - Estymacja punktowa vs. estymacja przedziałowa: przedział ufności, poziom istotności, wartość p
   - Testowanie hipotez dotyczących średniej i wariancji w jednej próbie

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

Najbliższe terminy
2026-12-14 (Zdalnie)
2026-12-14 (Warszawa)

- — Testowanie hipotez dotyczących porównania średniej i wariancji w dwóch próbach zależnych i dwóch próbach niezależnych
- — Analiza korelacji liniowej Pearsona i Spearmana dla zmiennych ilościowych
- — Analiza zależności dla zmiennych jakościowych (test chi-kwadrat)
- — Testowanie hipotez dotyczących średniej i wariancji w więcej niż dwóch próbach niezależnych (jednoczynnikowa ANOVA), testy porównań wielokrotnych
- — Testy nieparametryczne do porównywania rozkładów: test Wilcoxona-Manna-Whitney'a, test Kruskala-Wallisa

4. **Analiza regresji liniowej w R**
   - — Założenia modelu regresji liniowej, interpretacja wyników regresji (parametry, dopasowanie, istotność zmiennych i modelu)
   - — Jakościowe zmienne objaśniające w modelu
   - — Wykrywanie obserwacji nietypowych
   - — Metody doboru zmiennych
   - — Diagnostyka modelu, testowanie spełnienia założeń: testowanie normalności reszt, testowanie homoskedastyczności, testowanie braku autokorelacji
   - — Postępowanie w przypadku niespełnienia założeń
   - — Regresja logistyczna
   - — Prognozowanie na podstawie stworzonego modelu, sprawdzanie jakości prognozy, podział na próbę uczącą i testową
5. **Wizualizacja danych z wykorzystaniem pakietu ggplot2**
   - — Przegląd typów wykresów
   - — Dodawanie i dostosowywanie elementów wykresu
   - — Wizualizacja danych w grupach i podgrupach
   - — Tworzenie wykresów z wykorzystaniem kreatora (pakiet esquisse)
   - — Wykresy interaktywne (pakiet ggplotly)
   - — Przykładowe pakiety wspomagające tworzenie wykresów
   - — Zestawianie kilku wykresów w jednym oknie graficznym
   - — Eksport wykresów do różnych formatów plików

Przeznaczenie i wymagania

Znajomość statystyki ani R nie jest konieczna do uczestnictwa w szkoleniu.

Certyfikaty

Uczestnicy szkolenia otrzymują imienne certyfikaty sygnowane przez ALX.

Lokalizacje

- — Warszawa – ul. Jasna 14/16A
- — Zdalnie – zajęcia realizowane poprzez platformę Zoom
- — Kraków – ul. św. Filipa 23
- — Warsaw (English) – Jasna 14/16A
- — Online (English) – your home, office or wherever you want
- — na życzenie dowolne miejsce w Polsce, lub UE (zajęcia prowadzone w języku angielskim)

Cena szkolenia

1990 PLN netto (VAT 23%)

W cenę szkoleń organizowanych w naszej siedzibie wliczone są:

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

Najbliższe terminy
2026-12-14 (Zdalnie)
2026-12-14 (Warszawa)

— autorskie materiały szkoleniowe,
— indywidualne stanowisko komputerowe do pracy podczas zajęć,
— certyfikaty ukończenia szkolenia,
— drobny poczęstunek oraz ciepłe i zimne napoje,
— możliwość jednorazowego kontaktu z instruktorem (instruktorami) po szkoleniu i zadawania pytań dotyczących materiału szkolenia.

Cena szkolenia nie zawiera obiadów. Można je dokupić w cenie 35 zł netto za obiad.

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

Najbliższe terminy
2026-12-14 (Zdalnie)
2026-12-14 (Warszawa)