

Analiza danych w języku R

Opis

Szkolenie Analiza danych w R jest dla wszystkich tych, którzy chcą poznać tajniki analizy danych w zyskującym popularność języku statystycznym R i RStudio. Szkolenie prowadzone jest w formie warsztatów i ćwiczeń, gdzie teoria przeplata się z praktycznymi przykładami, z którymi możesz spotkać się w pracy zawodowej. Kurs prowadzony jest od podstaw po elementy zaawansowane.

Czego Cię nauczymy

Język R

Język R jest interpretowanym językiem programowania stosowanym do obliczeń statystycznych i wizualizacji wyników. R jest podstawowym językiem programowania w bioinformatyce, oraz jest wykorzystywany w wielu firmach, w tym m.in. Facebook, Google, Merck, Altera, Pfizer, LinkedIn, Shell, Novartis, Ford, Mozilla czy Twitter. Producenci komercyjnych pakietów statystycznych (SPSS, SAS, RapidMiner, Statistica) oferują dedykowane mechanizmy zapewniające ich współpracę z R.

RStudio

RStudio to zintegrowane środowisko programistyczne dla języka R, języka programowania do obliczeń statystycznych i grafiki. Na szkoleniu nauczysz się pracować z narzędziem RStudio, tworzyć skrypty, wizualizacje oraz wykresy.

Program szkolenia

1. Wprowadzenie do języka R

- Specyfika języka R
- Instalacja R i RStudio
- Praca z konsolą
- Środowisko pracy
- Konfiguracja RStudio
- Dobre praktyki pisania kodu
- Korzystanie z pomocy
- Instalacja pakietów
- Dokumentacja pakietów

2. Typy danych, zmienne i operatory

- Przypisywanie zmiennych
- Zmienne liczbowe
- Zmienne tekstowe
- Zmienne logiczne
- Zmienne typu Data
- Wektory typu factor
- Operatory matematyczne, logiczne i ciągów znaków
- Funkcje wbudowane (liczbowe, tekstowe, daty)
- Funkcje konwertujące

3. Praca z wektorami

- Czym jest wektor
- Rodzaje wektorów
- Operacja na wektorach
- Tworzenie i praca z wektorami
- Operacja na liczbach, datach, tekstach
- Odwołania do wektorów
- Filtrowanie wektorów
- Generowanie wektorów liczb losowych

4. Praca z obiektami

- Wielowymiarowe obiekty
- Tablice, macierze, listy, ramki danych
- Import i eksport danych

5. Instrukcje sterujące i programowanie

- Instrukcje warunkowe
- Pętle sterujące
- Funkcje do pracy z danymi

6. Metody statystyczne

- Wstęp do metod statystycznych
- Statystyka opisowa
- Rozkład zmiennej ilościowej
- Rozkład zmiennej jakościowej
- Mediana, dominanta, kwartyl, decyl, kwantyl
- Tworzenie własnych funkcji statystycznych
- Tabele kontyngencji
- Analiza korelacji
- Modele liniowe

7. Przetwarzanie danych

- Ramka danych
- Przeglądanie danych
- Filtrowanie i wybieranie danych
- Sortowanie danych
- Agregacja danych
- Tworzenie podsumowań
- Łączenie danych

8. Wizualizacja danych

- Wykresy jednej zmiennej (punktowe, kolumnowe, słupkowe, kołowe itp.)
- Tworzenie histogramów
- Wykresy wielu zmiennych
- Wiele wykresów obok siebie
- Elementy wykresu (tytuły, etykiety, osie, legenda, słupki błędów)
- Formatowanie wykresów
- Zapisywanie wykresów
- Wykorzystanie motywów i palet kolorów

Czas trwania

3 dni | 24 godziny zajęć

Certyfikat

Uczestnicy szkolenia otrzymują imienne certyfikaty sygnowane przez Expose Sp. z o.o.

Cena szkolenia

1 990 PLN netto (VAT 23%) za osobę (szkolenie grupowe)

12 990 PLN netto (VAT 23%) w trybie indywidualnym (1 na 1 z trenerem)

Cena szkolenia zawiera

- ✓ zapewnienie autorskich materiałów szkoleniowych dla uczestników szkolenia
- ✓ wystawienie certyfikatów po zakończonym szkoleniu
- ✓ rekomendacje dla uczestników szkolenia w zakresie dalszej pracy w obszarze szkolenia
- ✓ pakiet konsultacji z wykładowcą po zakończonym szkoleniu w razie jakichkolwiek niejasności przez okres 3 miesięcy
- ✓ całodzienny serwis kawowy oraz lunch