

Kod szkolenia: **R/BP**

Tytuł szkolenia: **Warsztaty analityka danych w R**

Dni: **3**

## Opis:

### Adresaci szkolenia

Szkolenie przeznaczone jest przede wszystkim dla analityków danych, którzy mają już pewne doświadczenie i chcą dowiedzieć się "jak to zrobić w R". Jednak podejście od podstaw umożliwi także wstępne zapoznanie się ze specyfiką tej dziedziny wiedzy osobom, które dopiero stawiają swoje pierwsze kroki.

### Cel szkolenia

Uczestnicy szkolenia zapoznają się z językiem R, a także szeroką gamą pakietów, które towarzyszą temu środowisku i pozwalają w wygodny i szybki sposób dokonać analizy danych. Szkolenie pozwoli na łączenie wielu narzędzi, co z kolei umożliwi wykonywanie bardziej skomplikowanych analiz z użyciem większego wachlarza technik.

### Wymagania

Szkolenie wymaga podstawowej umiejętności programowania w dowolnym języku lub podstaw programowania w języku R.

### Parametry szkolenia

3 \* 8 godzin (7 godzin netto) wykładów połączonych z warsztatami i ćwiczeniami.

### Program szkolenia:

1. Wstęp
  - Charakterystyka języka R
  - Porównanie z innymi językami
  - Dlaczego R?
2. Instalacja podstawowego środowiska
  - Instalacja R
  - Instalacja RStudio
  - Kontrola wersji z Git
    - Instalacja
    - Podstawowe operacje



## 3. RStudio

- Podstawowe okno pracy
- Historia komend
- Konsola R
- Kod źródłowy
- Korzystanie z plików pomocy

## 4. Typy danych i możliwości języka R niedostępne w innych językach

- Obliczenia zwektoryzowane
- Funkcje z rodziny apply
- Ramki danych
- Typ czynnikowy

## 5. Inżynieria danych:

- Wczytywanie i czyszczenie danych
- Przekształcanie danych przy użyciu bazowego R
  - Wybór poszczególnych kolumn
  - Wybór wierszy spełniających zadany warunek
  - Grupowanie danych
  - Liczenie podstawowych statystyk
  - Sortowanie
- Przekształcanie danych przy użyciu pakietu dplyr
  - Wybór poszczególnych kolumn
  - Wybór wierszy spełniających zadany warunek
  - Grupowanie danych
  - Liczenie podstawowych statystyk
  - Sortowanie

## 6. Wizualizacja

- Grafika niskopoziomowa
  - Barwy, punkty, odcinki, wielokąty
  - Zapisywanie wykresów w formatach grafiki rastrowej (np. jpg) i wektorowej (np. .pdf)
- Grafika wysokopoziomowa
  - Wykres słupkowy, kołowy
  - Wykres skrzynkowy, histogram
  - Wykres rozproszenia
- Pakiet ggplot2

## 7. Statystyka i uczenie maszynowe

- Gęstość, dystrybuanta, odwrotna dystrybuanta, losowanie z podanego rozkładu - przydatne funkcje w R
- Podstawowe testy statystyczne
- Regresja liniowa
- Regresja logistyczna
- Analiza skupień
- Drzewa decyzyjne

