

SZKOLENIE ŚREDNIO ZAAWANSOWANE

Warsztat analityka danych w R

R/BP

Czas trwania: 3 dni (24h)

Cele szkolenia

- Zapoznanie 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

Zalety

- Praktyka przed teorią - wszystkie szkolenia technologiczne prowadzone są w formie warsztatowej. Konieczna teoria jest wyjaśniana na przykładzie praktycznych zadań
- Konkretność umiejętności - w ramach każdego szkolenia rozwijamy praktyczne umiejętności związane z daną technologią i tematyką
- Nauka z praktykami - wszyscy trenerzy na co dzień pracują w projektach, gwarantuje to dostęp do eksperckiej wiedzy i praktycznego know-how

Dla kogo?

- Analitycy danych, którzy mają już pewne doświadczenie i chcą dowiedzieć się jak to zrobić w R”
- Osoby pracujące w działach sprzedaży, marketingu, finansów lub analiz, które chcą rozwijać kompetencje związane z analizą danych, wykorzystując rozwiązania wykraczające poza narzędzia typu excel
- Osoby z zapałem do programowania i znający już pracę analityka, jako, że szkolenie skupia się na narzędziach a nie samych technikach

Wymagania

- Znajomość dowolnego języka programowania nie jest wymagana, ale może znacząco ułatwić przyswajanie materiału

Program

1. Wstęp
 - a. Charakterystyka języka R
 - b. Porównanie z innymi językami
 - c. Dlaczego R?



2. Instalacja podstawowego środowiska

- a. Instalacja R
- b. Instalacja RStudio
- c. Kontrola wersji z Git
 - Instalacja
 - Podstawowe operacje

3. RStudio

- a. Podstawowe okno pracy
- b. Historia komend
- c. Konsola R
- d. Kod źródłowy
- e. Korzystanie z plików pomocy

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

- a. Obliczenia zwektoryzowane
- b. Funkcje z rodziny apply
- c. Ramki danych
- d. Typ czynnikowy

5. Inżynieria danych

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

6. Wizualizacja

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

7. Statystyka i uczenie maszynowe

- a. Gęstość, dystrybuanta, odwrotna dystrybuanta, losowanie z podanego rozkładu - przydatne funkcje w R
- b. Podstawowe testy statystyczne



- c. Regresja liniowa
- d. Regresja logistyczna
- e. Analiza skupień
- f. Drzewa decyzyjne

