

Program szkolenia

Praktyczna analiza danych w R

Prowadzący: dr inż. Artur Suchwałko

1. Podstawy praktycznej analizy danych: przygotowanie danych i analiza opisowa
 - rodzaje cech statystycznych – cechy jakościowe i ilościowe
 - przygotowanie danych do analiz (wybór podzbiorów, standaryzacja, proste przekształcenia)
 - grupowanie danych (tablice wieloznaczne, tabele przestawne)
 - analiza opisowa danych – podstawowe wskaźniki i ich własności
 - grafika w analizie opisowej – użyteczne wykresy i ich interpretacja (histogramy, wykresy słupkowe i kołowe, wykres pudełkowy, wykresy rozrzutu i inne)
 - wizualizacja danych kategoriycznych
 - problem jakości danych: obserwacje brakujące i nietypowe
 - przykłady analizy opisowej dla danych rzeczywistych
2. Podstawy statystyki, rozkłady zmiennych losowych oraz próbkowanie danych
 - czym jest statystyka?
 - możliwe cele i rodzaje badań statystycznych
 - etapy w badaniu statystycznym
 - podstawy modeli statystycznych – zmienne losowe dyskretne i ciągłe oraz ich rozkłady.
 - metody doboru próby
3. Elementy wnioskowania statystycznego
 - podstawy weryfikacji hipotez statystycznych – idea testowania hipotez, zastosowania praktyczne i przegląd najpopularniejszych testów statystycznych (testy istotności, zgodności i niezależności).
 - estymacja parametrów
 - przedziały ufności
 - testowanie hipotez statystycznych dla jednej populacji
 - testowanie hipotez statystycznych dla dwóch populacji
 - sprawdzanie hipotez statystycznych w praktyce
 - przykłady zastosowania metod wnioskowania statystycznego dla danych rzeczywistych
 - analiza wariancji
 - wybrane metody nieparametryczne
4. Model regresji liniowej

- analiza zależności dwóch zmiennych ilościowych
- ocena zależności zmiennych ilościowych: współczynnik korelacji i wykres rozrzutu
- prosty model regresji liniowej: założenia i ograniczenia modelu, interpretacja
- dopasowanie i diagnostyka modelu
- porównanie i wybór najlepszego modelu
- regresja wielokrotna
- wybór zmiennych do budowy modelu
- wykorzystanie dopasowanego modelu do prognozowania
- praktyczna analiza korelacji i regresji na przykładach danych rzeczywistych