

Statystyka opisowa w środowisku R

ZADANIA

Zad. 1. Baza danych *LakeHuron* zawiera informacje dotyczące corocznego poziomu wody w jeziorze Huron (podane w stopach) w latach 1875-1972.

- a) Zbuduj szereg przedziałowy dla analizowanych danych:
 - i. o równych długościach przedziałów,
 - ii. składający się z 5 klas o (w przybliżeniu) równych liczebnościach.
- b) W jakim przedziale określającym poziom wody w jeziorze znajduje się dominanta?

Zad. 2. Baza danych *Indometh* zawiera informacje na temat stężenia indometacyny (leku przeciwzapalnego) we krwi. Dokonaj analizy zmiennej `conc`, która zawiera dane liczbowe określające stężenie leku (mcg/ml).

- a) Jaka jest wielkość próbki?
- b) Z jakiego przedziału przyjmuje wartości analizowana zmienna?
- c) Skomentuj wzajemne położenie średniej i mediany.
- d) Zbadaj symetrię próbki.
- e) Jaka jest średnia obcięta z 50% środkowych obserwacji?
- f) Zbadaj koncentrację próbki.
- g) Jaka jest długość przedziału, w którym znajduje się 50% środkowych obserwacji?
- h) Ile wynosi poziom stężenia indometacyny, poniżej którego znajduje się 30% obserwacji?
- i) Ile wynosi poziom stężenia indometacyny, powyżej którego znajduje się 20% obserwacji?

Zad. 3 Baza danych *iris* zawiera dane dotyczące trzech gatunków irysów (gatunek kwiatu określa zmienna *Species*). Porównaj długości kielicha kwiatów irysa (zmienna *Sepal.Length*) w każdym z tych gatunków wykonując podstawową analizę statystyczną danych.