

Analiza skupień zestawu samochodów amerykańskich, europejskich i japońskich wyprodukowanych w latach 1970-1982

Jerzy Marczewski

Dawid Rozumkiewicz

Wojciech Pietraszuk

Karol Hetmański

Przemysław Chachaj

1 Streszczenie

2 Słowa kluczowe

3 Wprowadzenie

4 Przedmiot badania

4.1 Cel i zakres badania

4.2 (min. 1 cytowanie powiązane tematycznie i krótki opis co było badane)

4.3 Zmienne wybrane do analizy (opis i uzasadnienie zmien- nych oraz podział na stymulanty/destymulanty) mini- mum sześć zmiennych

Do analizy wybrano następujące cechy diagnostyczne:

- X_1 - mpg
- X_2 - liczba cylindrów
- X_3 - objętość skokowa silnika
- X_4 - hp
- X_5 - waga (podana w funtach)

- X_6 - przyspieszenie
- X_7 - model (rok modelowy auta wyrażony w dwóch ostatnich cyfrach roku)
- X_8 - kraj pochodzenia (1 - Stany Zjednoczone, 2 - Europa, 3 - Japonia)

4.4 Wstępna analiza danych

Statystyki opisowe

Wyniki średniej, mediany, minimum, maksimum, odchylenia standardowego, skośności dla każdej z cech zaokrąglone do dwóch miejsc po przecinku:

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
średnia	20.76	5.81	216.09	112.91	3139.32	15.31	74.35	1.46
mediana	19.2	6.0	225.0	100.0	3139.0	15.5	74.0	1.0
min	9	4	72	46	1613	9	70	1
max	43.1	8.0	455.0	230.0	5140.0	22.1	79.0	3.0
odchylenie standardowe	6.62	1.78	109.93	41.64	900.74	2.80	2.93	0.74
skośność	0.71	0.18	0.34	0.79	0.24	0.06	0.04	1.21

podstawowa wizualizacja np. boxplot, histogramy

braki danych, czy występują i jak je obsłużono

Braki danych występują jedynie dla cechy hp. Wszystkie samochody z brakiem danych zostały usunięte z danych za pomocą skryptu w R.

obserwacje odstające i w jaki sposób je obsłużono

5 Opis metod

5.1 wzory wraz z opisami oznaczeń

Metoda k-średnich

Metoda Warda

SOM

5.2 cytowanie pracy w której zaproponowano metodę/e- wentualnie pracy, w której użyto metodę

6 Rezultaty (w postaci tabelarycznej i/lub gra- ficznej oraz omówienie wyników)

7 Podsumowanie (ocena realizacji celu, odniesie- nie do pozycji z przeglądu literatury)

8 Bibliografia

- Algorithm AS 136: A K-Means Clustering Algorithm - J. A. Hartigan and M. A. Wong
- www.statystyka.az.pl
- www.wikipedia.org