

Komputerowa analiza danych

Zadanie 2

1. Cel

Implementacja i analiza działania klasycznego algorytmu k -średnich (ang. k -means), służącego do grupowania danych.

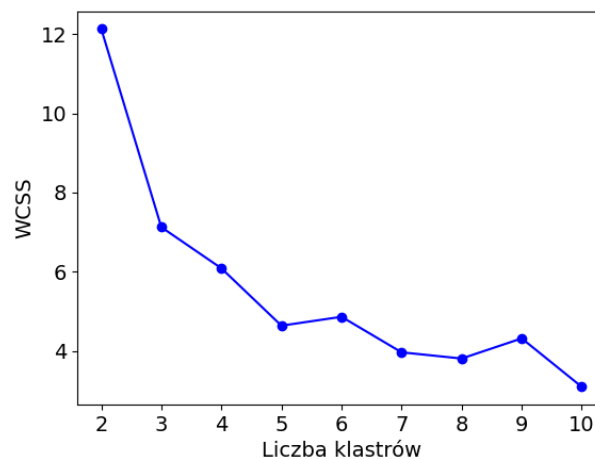
2. Wyniki

Zostały użyte następujące biblioteki języka Python:

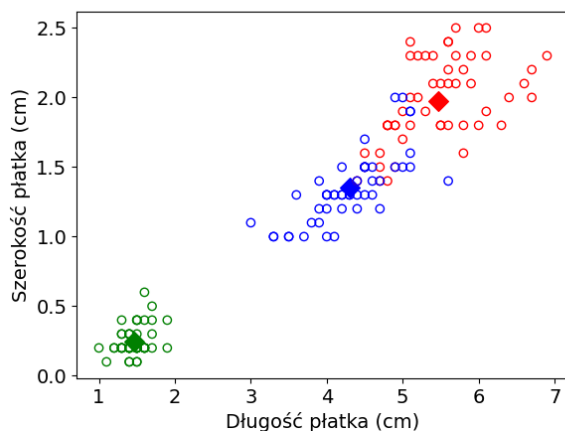
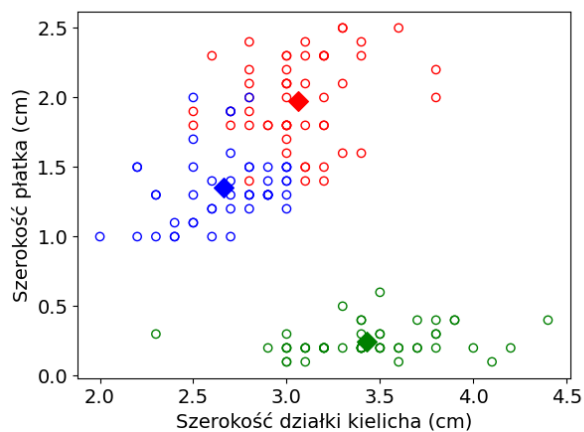
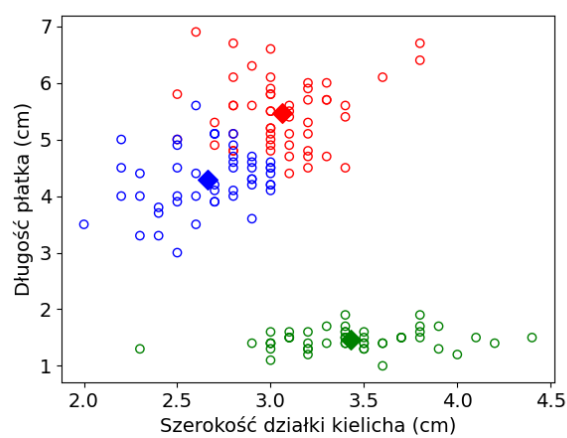
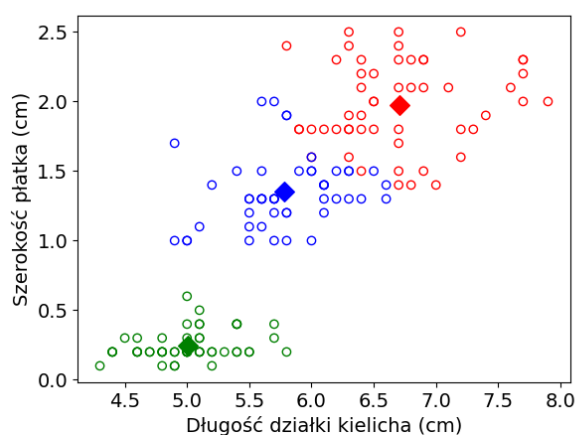
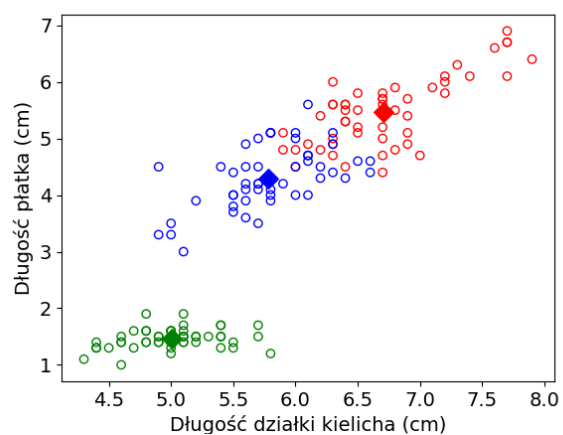
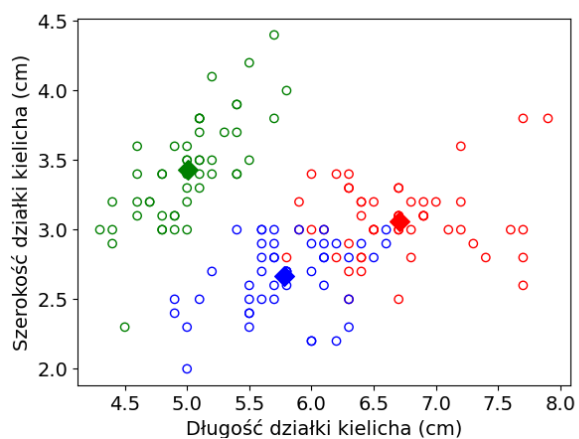
- Matplotlib do generowania wykresów,
 - Pandas do wczytania danych z pliku,
 - Numpy do niektórych obliczeń
- oraz napisany przez program realizujący algorytm k -średnich.

Tabela 1. Wartości WCSS oraz liczba iteracji algorytmu k -średnich w zależności od ilości klastrów

k	Liczba iteracji	WCSS
2	6	12,12
3	5	7,12
4	7	6,08
5	7	4,63
6	7	4,85
7	7	3,96
8	7	3,80
9	8	4,31
10	6	3,09



Rysunek 1. Wykres przedstawiający wartości WCSS w zależności od ilości klastrów



Rysunek 2. Wykresy dla każdej z sześciu możliwych par cech