

Komputerowa Analiza Danych

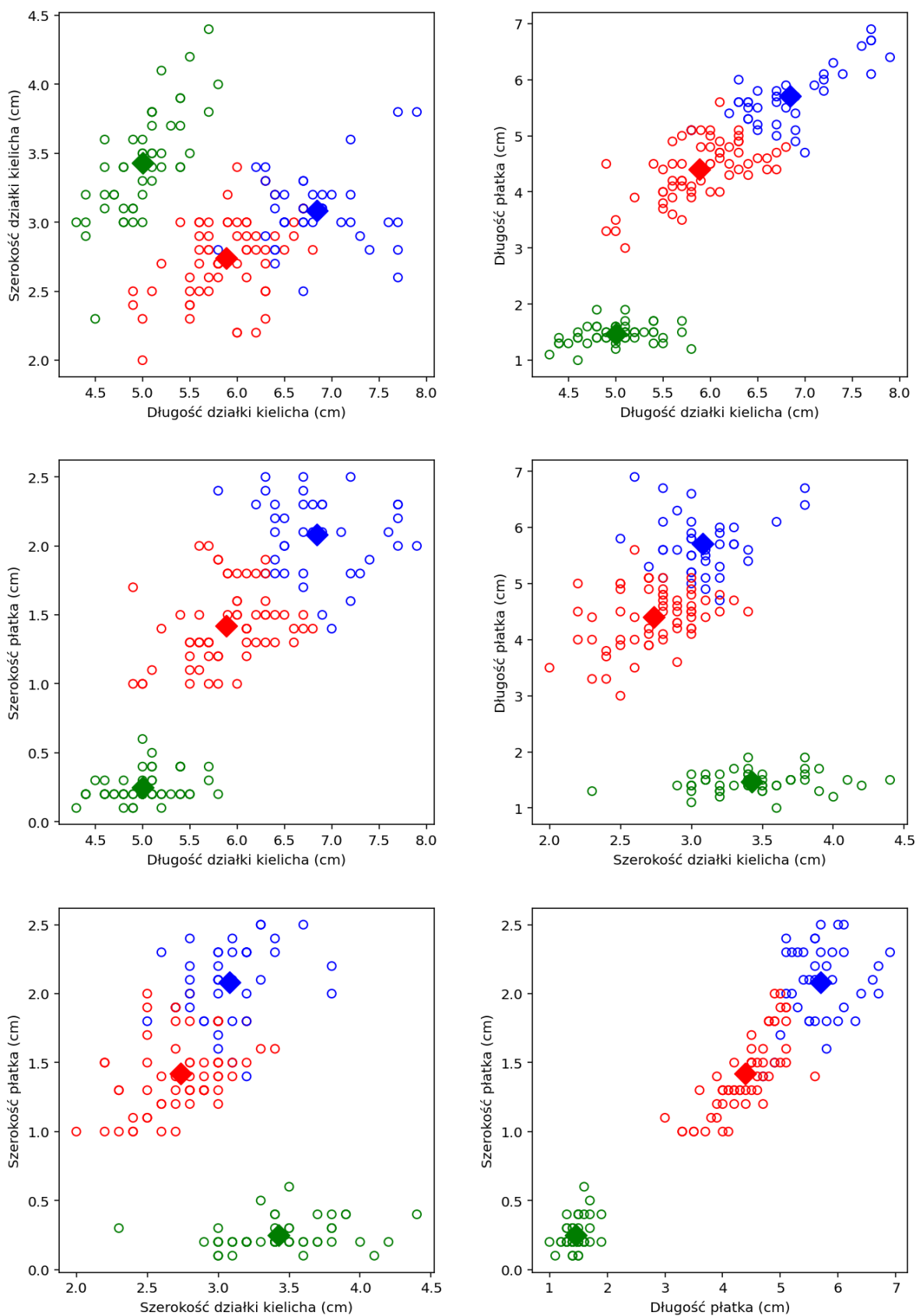
Zadanie 2

1. Cel

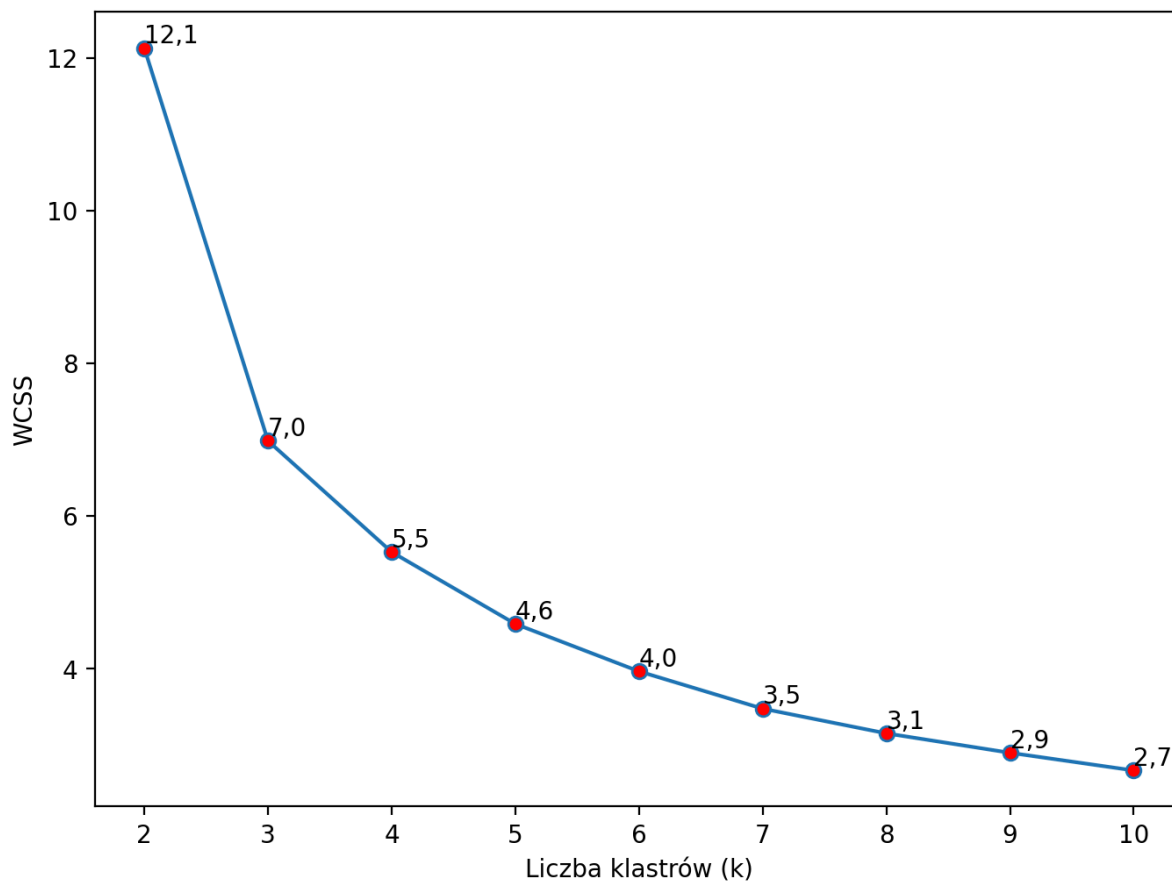
Implementacja i analiza działania klasycznego algorytmu k -średnich (ang. *k-means*), służącego do grupowania danych.

2. Wyniki

Poniższe wyniki zostały uzyskane za pomocą kodu napisanego w języku Python. Algorytm k -średnich został zaimplementowany własnoręcznie. Do stworzenia wykresów został zaimplementowany skrypt Python używający biblioteki Matplotlib.



Rysunek 1: Końcowa lokalizacja centrów w przestrzeni cech, po zakończeniu algorytmu k-średnich. Przedstawiono dodatkowo przynależność poszczególnych obiektów do centrów. Algorytm działał przez 100 cykli.



Rysunek 2: Wykres przedstawiający wartości $WCSS$ dla kolejnych klastrow k

Tabela 1: Porównanie najlepszych *within-cluster sum of squares*, wybranych z 10 prób

k	najniższa $WCSS$	liczba iteracji
2	12,1	100
3	7,0	100
4	5,5	100
5	4,6	100
6	4,0	100
7	3,5	100
8	3,1	100
9	2,9	100
10	2,7	100