

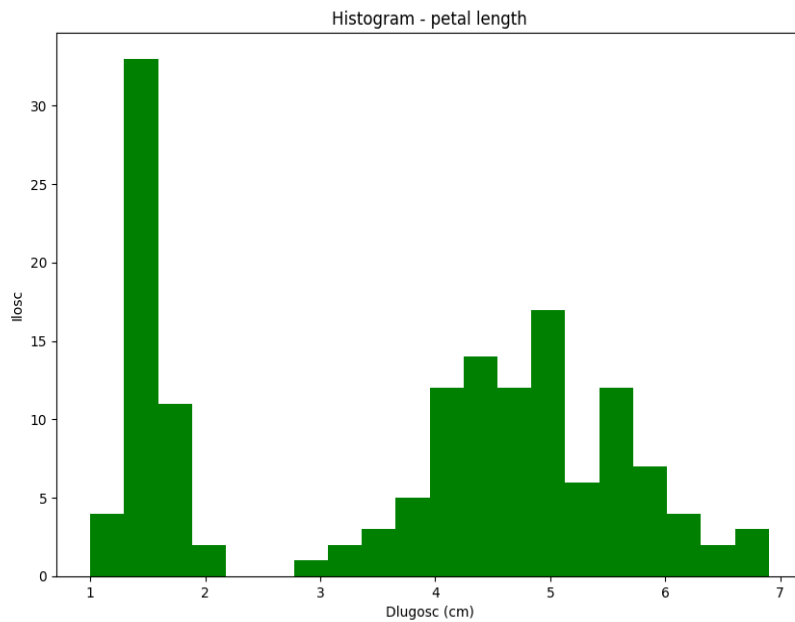
Grupowanie danych Iris metodą k-średnich

Szymon Prud, Informatyka 3 semestr

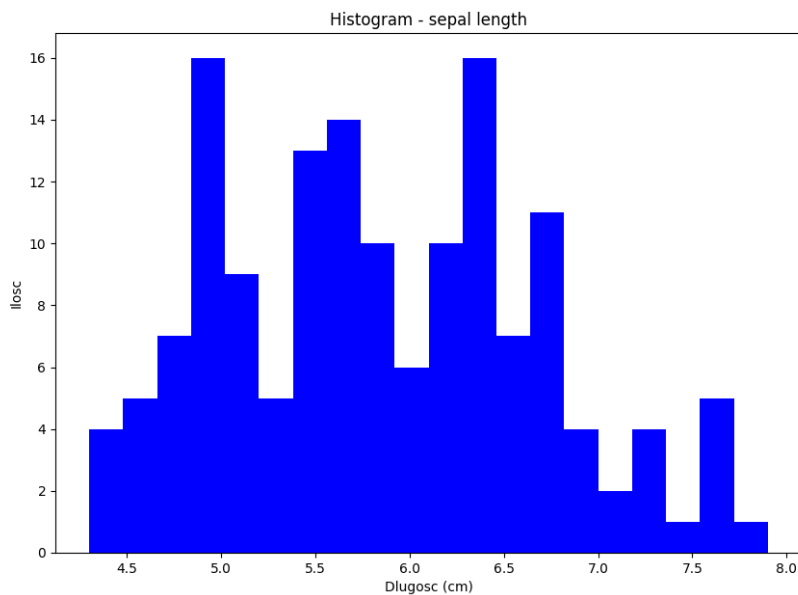
Numer indeksu: 394181

1. Histogramy

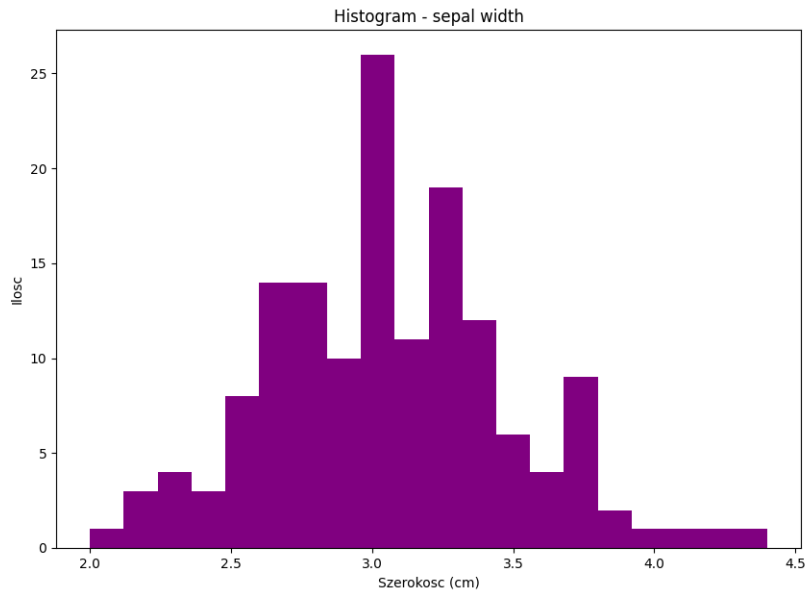
-Petal length



-Sepal length



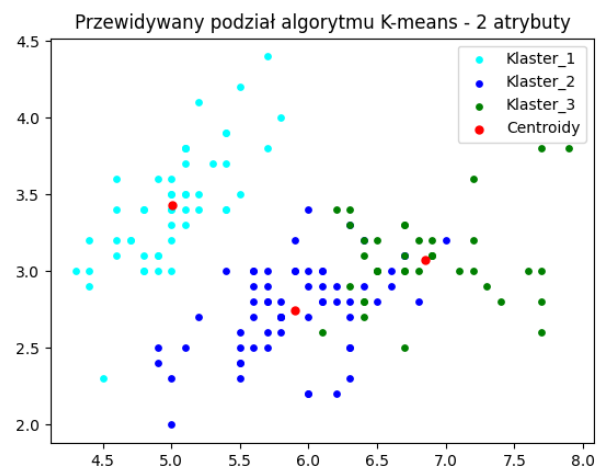
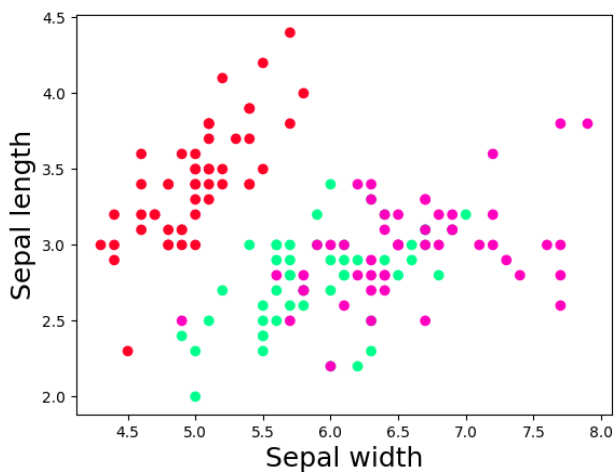
-Sepal width



2. Klastrowanie wzorców metodą k-means dla 2 atrybutów (2D):

-Wizualizacja bazy danych przed posortowaniem:

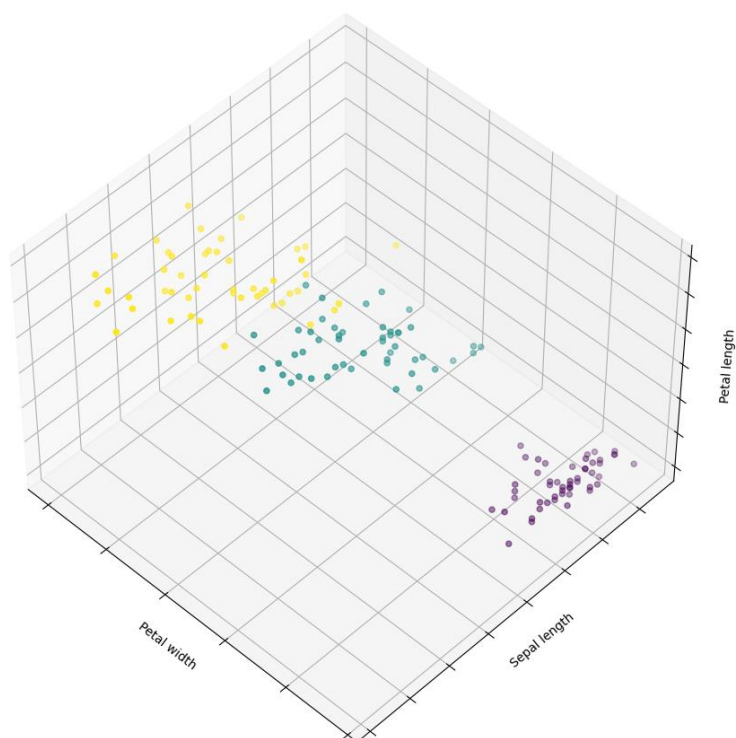
-Wizualizacja danych po posortowaniu k-means:



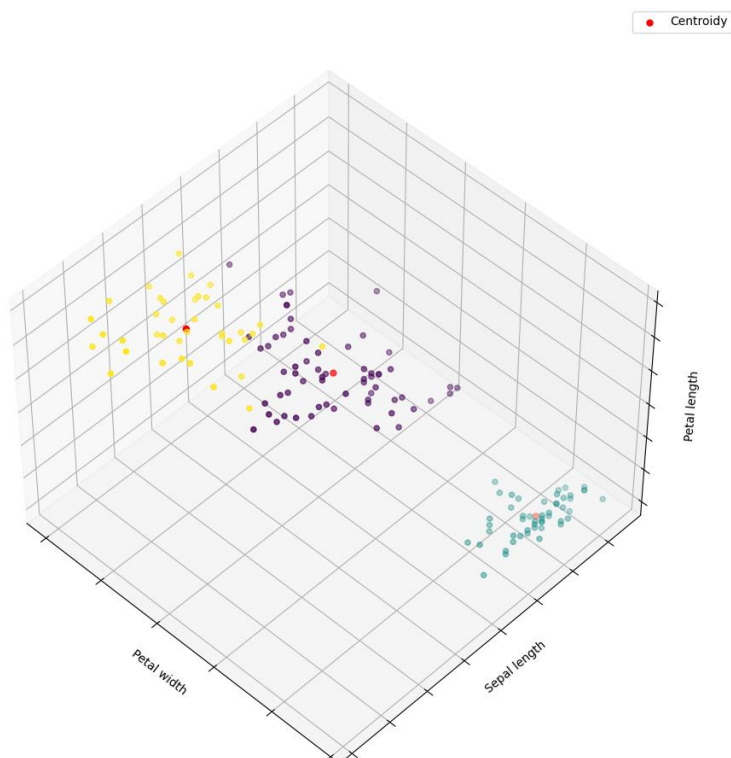
Dla dwóch atrybutów po sortowaniu do innych grup trafiło 13 wzorców (ok. 8%).

3. Klastrowanie wzorców metodą k-means dla 3 wzorców (3D):

- Wizualizacja przed posortowaniem:



- Wizualizacja po posortowaniu:



Dla trzech atrybutów do innych grup trafiło 14 wzorców (około 9%)

4. Klastrowanie wzorców metodą k-means dla 4 wzorców (4D):

W taki sposób wyglądają współrzędne centroidów utworzone przez algorytm k-means:

```
[[5.9016129 2.7483871 4.39354839 1.43387097]  
 [5.006      3.428      1.462      0.246      ]  
 [6.85       3.07368421 5.74210526 2.07105263]]
```

Dokładność sortowania wynosi w tym przypadku ok. 89%

Możemy więc wywnioskować, że wraz ze wzrostem ilości atrybutów spada skuteczność algorytmu.

Kod na podstawie którego zostały wygenerowane powyższe dane został załączony w pliku.