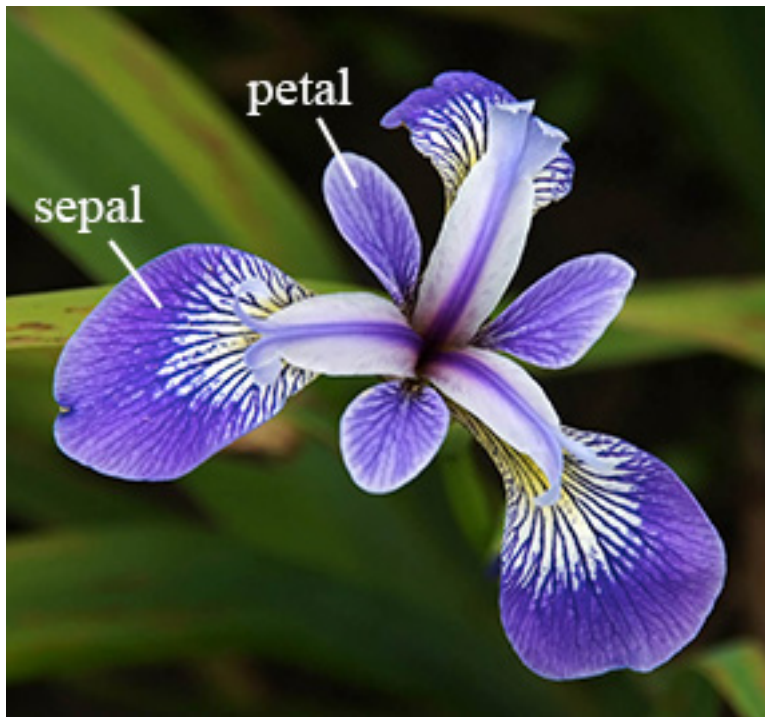

PPPD - Lab. 08

Copyright ©2021 M. Śleszyńska-Nowak i in.

Zbiór danych `iris` oryginalnie pochodzi z artykułu naukowego [Fisher R. A., The use of multiple measurements in taxonomic problems, *Annals of Eugenics* **7**, 1936, s. 179–188] i składa się z wyników pomiarów czterech następujących charakterystyk 150 badanych kosaćców (irysów):

- kolumny 1 i 2: długość i szerokość kielicha (sepal length, sepal width);
- kolumny 3 i 4: długość i szerokość płatków (petal length, petal width).



Rysunek 1: „Składowe” kosaćców (irysów)

Załadujesz go, wywołując następujący kod:

```
import csv
iris = []
f = open("iris.csv", "r") # r=do odczytu
for row in csv.reader(f):
    for i in range(len(row)):
        row[i] = float(row[i]) # konwersja z str na float
    list.append(iris, row) # == A.append(row)
f.close()
```

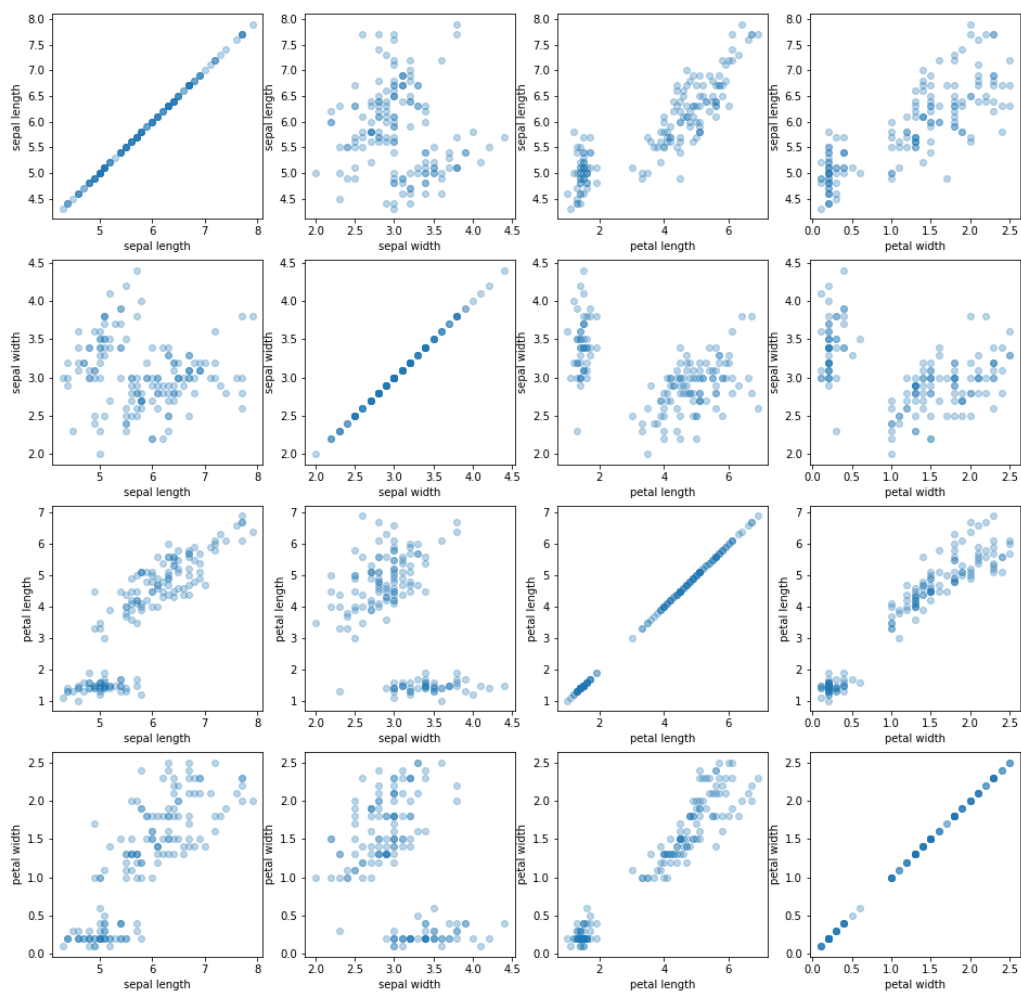
Od teraz `iris` jest macierzą typu 150×4 .

-
1. Dokonaj transpozycji ww. zbioru. Taka postać będzie łatwiejsza w dalszej obróbce, bowiem `iris[i]` będzie oznaczało *i*-tą kolumnę (będzie to lista o 150 elementach).
 2. Narysuj wykresy rozproszenia dla każdej pary zmiennych (por. rysunek poniżej). Jeden wykres rozproszenia narysujesz, wywołując funkcję `plt.scatter(zmienna_x, zmienna_y)`. Podzielisz jeden wykres na podwykresy, wywołując `plt.subplot(nrows, ncols, plot_number)`, zob. https://matplotlib.org/devdocs/api/_as_gen/matplotlib.pyplot.subplots.html Etykiety osi ustawisz, wywołując `plt.xlabel(napis)` lub `plt.ylabel(napis)`.

Kod generujący wykresy zarzyj wewnątrz następującego bloku instrukcji:

```
import matplotlib.pyplot as plt
plt.figure(figsize=[16,16], dpi=72)
#...
plt.savefig("output.png")
```

3. Oblicz wartość współczynnika korelacji liniowej Pearsona między każdą parą zmiennych i zapisz ją w postaci macierzy 4×4 . Zwróć uwagę, że macierz ta będzie symetryczna i będzie miała jedynki na głównej przekątnej.
4. Dla każdego podwykresu dopasuj proste metodą najmniejszych kwadratów i narysuj je (jak w jednym z zadań serii nr 3).



Rysunek 2: Wykresy rozproszenia dla każdej pary zmiennych