

Analiza zbioru Iris



Cel analizy

Czy na podstawie pomiarów poszczególnych części kwiatu można znaleźć różnice pomiędzy trzema gatunkami irysów?

klasa	dlugosc_kielicha	szerokosc_kielicha	dlugosc_platka	szerokosc_platka
Iris-setosa	5.000000	3.500000	1.300000	0.300000
Iris-virginica	6.400000	2.800000	5.600000	2.100000
Iris-versicolor	5.900000	3.000000	4.200000	1.500000
Iris-versicolor	6.000000	3.400000	4.500000	1.600000
Iris-versicolor	6.700000	3.100000	4.400000	1.400000

Krok 1: Ogólny przegląd danych

Dane zawierają 150 wierszy oraz 5 kolumn.

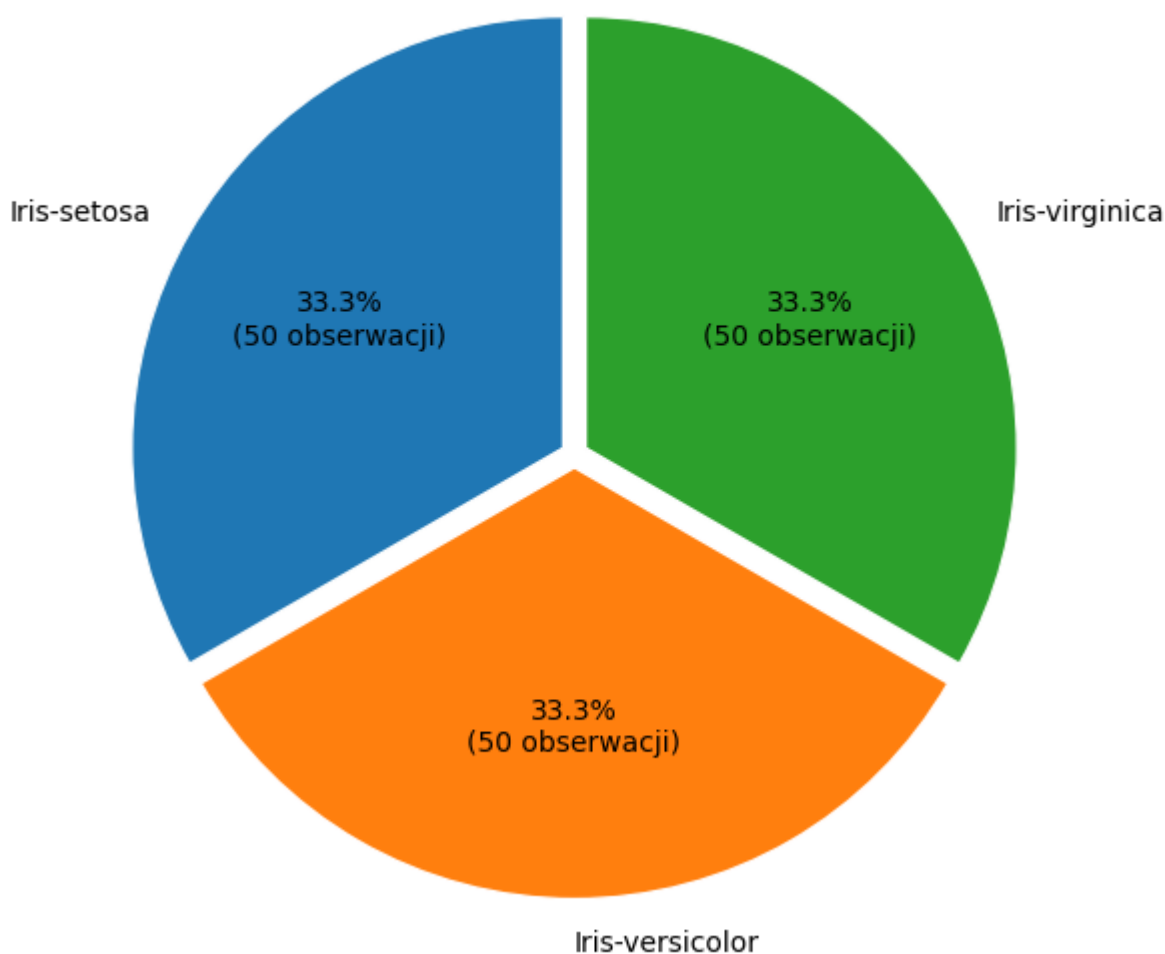
Dane dotyczą 3 gatunków irysów, konkretnie Iris-setosa, Iris-versicolor, Iris-virginica.

W naszych danych każdy z gatunków można rozróżnić również według 4 unikatowych cech, czyli: dlugosc_kielicha, szerokosc_kielicha, dlugosc_platka, szerokosc_platka.

Dane są wręcz idealnie zbalansowane, każdy z gatunków posiada po 50 rekordów.

klasa	dlugosc_kielicha	szerokosc_kielicha	dlugosc_platka	szerokosc_platka
Iris-setosa	50	50	50	50
Iris-versicolor	50	50	50	50
Iris-virginica	50	50	50	50

Udział gatunków w zbiorze Iris



Spostrzeżenia:

- Mamy łącznie 150 rekordów.

- Są one podzielone na 3 gatunki irysów: 'Iris-setosa', 'Iris-versicolor', 'Iris-virginica'.
- Wszystkie 3 zawierają równo po 50 obserwacji.
- Każdy z gatunków rozróżniamy również poprzez podane wartości opisujące długość i szerokość kielicha, jak i jego płatków.

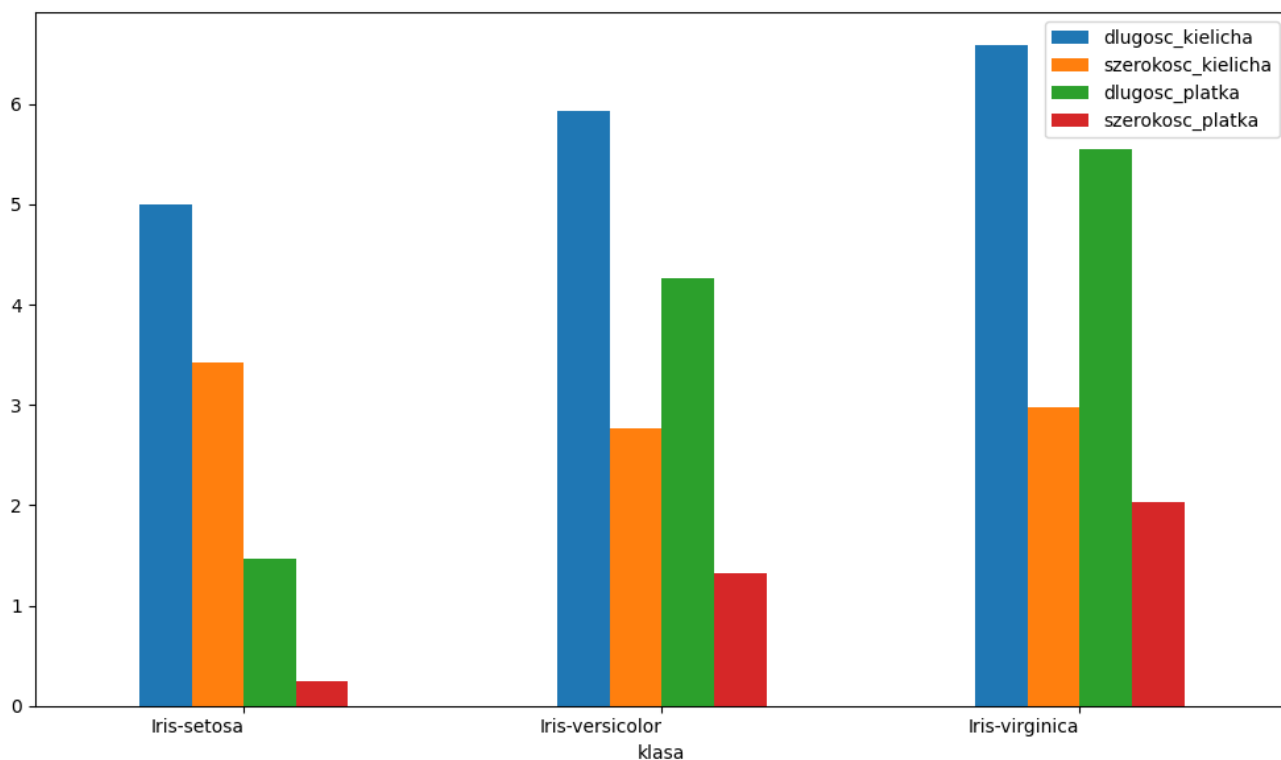
Krok 2: Brakujące dane

W danych dotyczących Irysów całkowita liczba brakujących danych wynosi 0.

Krok 3: Analiza pojedynczych kolumn

klasa	dlugosc_kielicha	szerokosc_kielicha
Iris-setosa	5.800000	4.400000
Iris-versicolor	7.000000	3.400000
Iris-virginica	7.900000	3.800000

klasa	dlugosc_platka	szerokosc_platka
Iris-setosa	1.900000	0.600000
Iris-versicolor	5.100000	1.800000
Iris-virginica	6.900000	2.500000



Spostrzeżenia:

1. *Iris setosa*

- Rozmiar kielicha: długość ~5.0, szerokość ~3.4
- Rozmiar płatków: długość ~1.46, szerokość ~0.24

3. *Iris virginica*

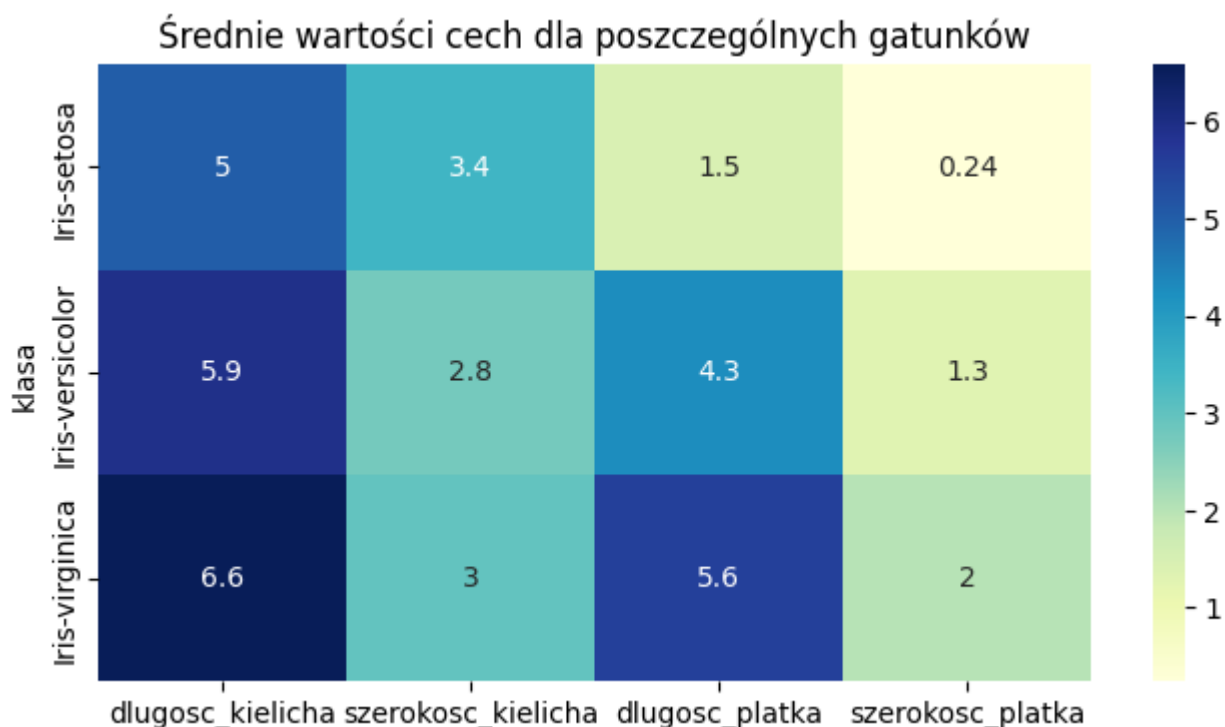
- Rozmiar kielicha: długość ~6.58, szerokość ~2.97
- Rozmiar płatków: długość ~5.55, szerokość ~2.02

2. *Iris versicolor*

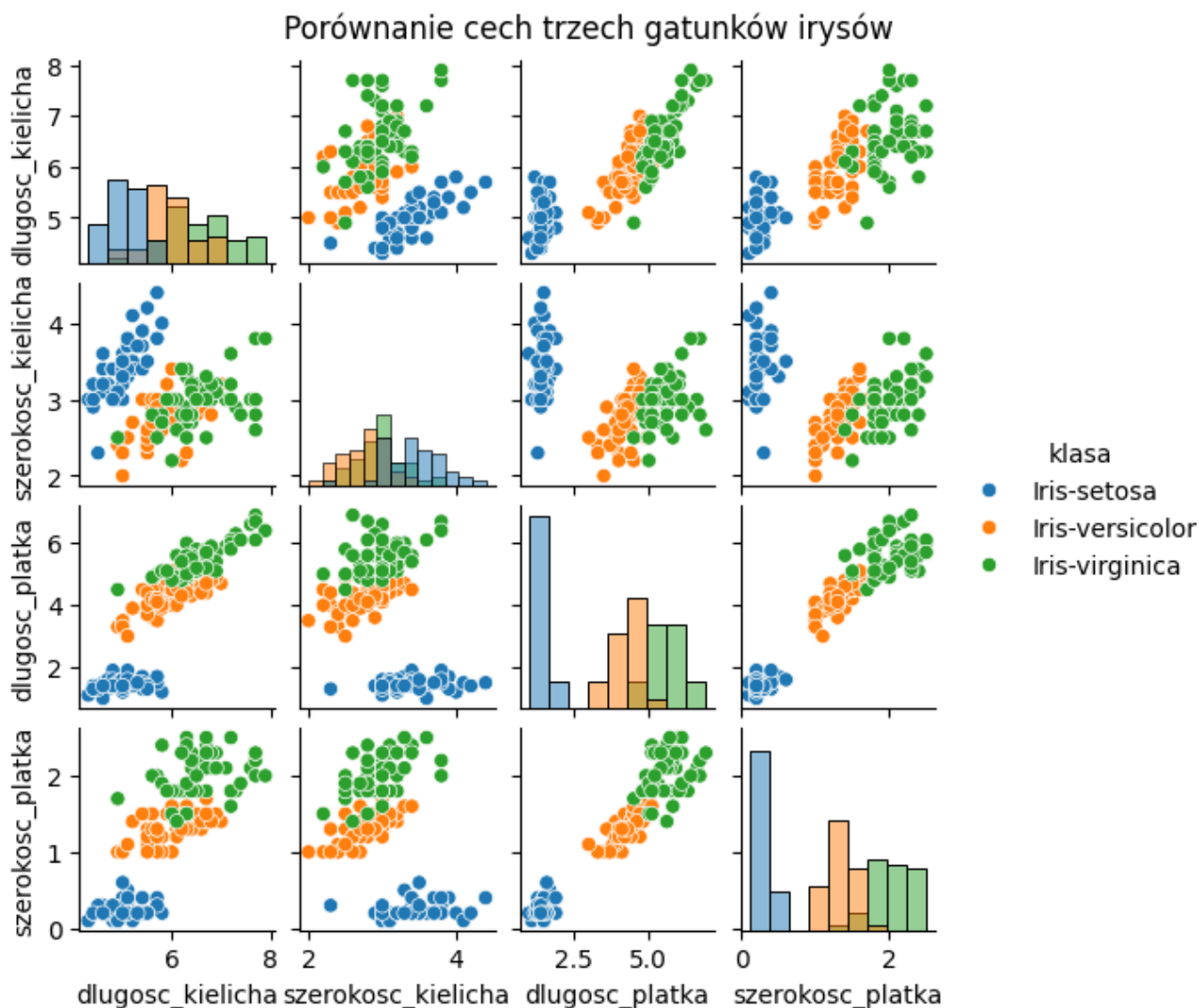
- Rozmiar kielicha: długość ~5.93, szerokość ~2.77
- Rozmiar płatków: długość ~4.26, szerokość ~1.32

Na pierwszy rzut oka da się dostrzec, że pod względem wielkości płatków dominuje odmiana *Iris virginica*, na drugim miejscu znajduje się *Iris versicolor*, a ostatni *Iris setosa*. Jeśli chodzi o rozmiary kielicha, z długością jest analogicznie jak w przypadku płatków, natomiast pod względem szerokości różnice nie są duże, z drobną przewagą na korzyść odmiany *Iris setosa*.

Krok 4: Porównanie cech



Im ciemniejszy kolor, tym większa średnia wartość danej cechy.



Na wykresie widać, że obserwacje *Iris versicolor* i *Iris virginica* częściowo się nakładają, natomiast *Iris setosa* tworzy wyraźnie odrębną grupę.

Spostrzeżenia:

1. *Iris setosa*

Najsilniejsze korelacje:

- Długość_kielicha, szerokość kielicha - całkiem silna dodatnia korelacja ~ 0.747
- Długość płatka, szerokość płatka - dość słaba dodatnia korelacja ~ 0.306

2. *Iris versicolor*

Najsilniejsze korelacje:

- Długość_kielicha, szerokość kielicha - silna dodatnia korelacja ~ 0.786
- Długość płatka, szerokość płatka - również silna dodatnia korelacja ~ 0.754

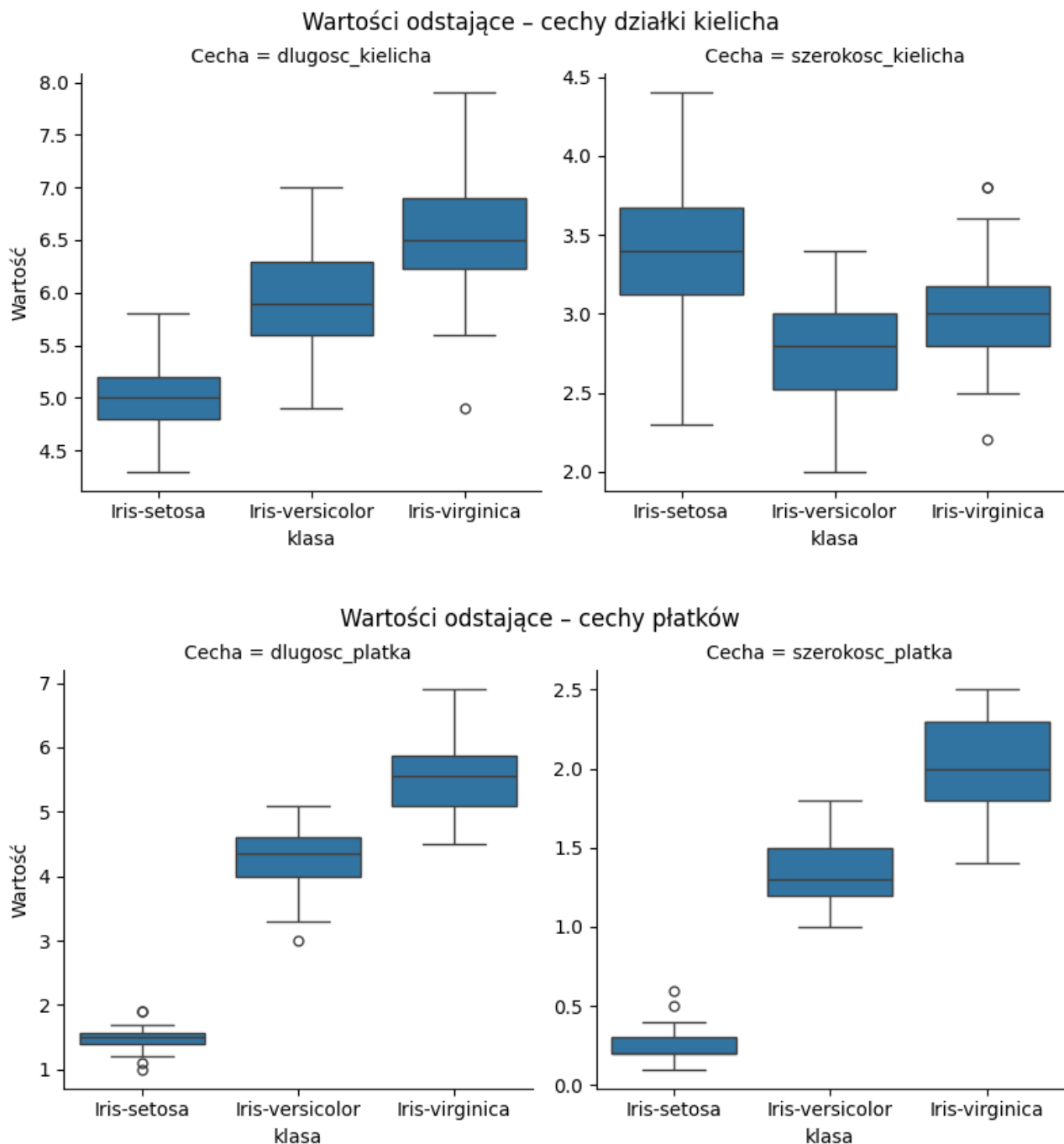
3. *Iris virginica*

Najsilniejsze korelacje:

- Długość_kielicha, szerokość kielicha - bardzo silna dodatnia korelacja ~ 0.864
- Długość płatków, szerokość płatków - średnia korelacja ~ 0.537

We wszystkich gatunkach dodatnia korelacja długości i szerokości płatków sugeruje, że większe płatki są zazwyczaj również szersze.

Krok 5: Wartości odstające



Spostrzeżenia:

W większości cech nie występują liczne wartości odstające. Pojedyncze obserwacje znajdujące się poza zakresem wąsów nie stanowią istotnego problemu i mogą wynikać z naturalnej zmienności

biologicznej badanych roślin.

Najważniejsze wnioski

- zbiór zawiera 150 obserwacji,
- dane są idealnie zbalansowane (50 obserwacji na gatunek),
- nie występują brakujące dane,
- największe różnice między gatunkami dotyczą płatków,
- *Iris setosa* jest wyraźnie oddzielona od pozostałych,
- największe wartości cech osiąga *Iris virginica*.

***Iris setosa*:**

Zauważalnie najmniejszy z gatunków, chociaż posiada najszerszy kielich wśród całej grupy Irysów z tej analizy. Jego rozmiary to:

- Rozmiar kielicha: długość ~5.0cm, szerokość ~3.4cm
- Rozmiar płatków: długość ~1.46cm, szerokość ~0.24cm

***Iris versicolor*:**

Pod względem wielkości większości cech zajmuje pozycję pośrednią pomiędzy *Iris setosa* a *Iris virginica*. Jego rozmiary to:

- Rozmiar kielicha: długość ~5.93cm, szerokość ~2.77cm
- Rozmiar płatków: długość ~4.26cm, szerokość ~1.32cm

***Iris virginica*:**

Zdecydowanie największy pomiędzy wszystkimi trzema gatunkami, w szczególności dominuje pod względem wielkości płatków - ich rozmiary wyraźnie przewyższają jego konkurentów. Poszczególne rozmiary to:

- Rozmiar kielicha: długość ~6.58cm, szerokość ~2.97cm
- Rozmiar płatków: długość ~5.55cm, szerokość ~2.02cm

0

0