

Cel analizy

Celem tej analizy było sprawdzenie, **które cechy fizyczne kwiatów Iris** pozwalają najlepiej rozróżnić trzy gatunki:

- *Iris setosa*
- *Iris versicolor*
- *Iris virginica*

Dataset zawiera **150 obserwacji** oraz **4 zmienne opisujące kształt kwiatu**:

Cecha	Opis
Sepal length	Długość kielicha
Sepal width	Szerokość kielicha
Petal length	Długość płatka
Petal width	Szerokość płatka

Badanie miało na celu **zrozumienie różnic między gatunkami** oraz ocenę, czy dane są stabilne i mogą zostać wykorzystane do modelowania.

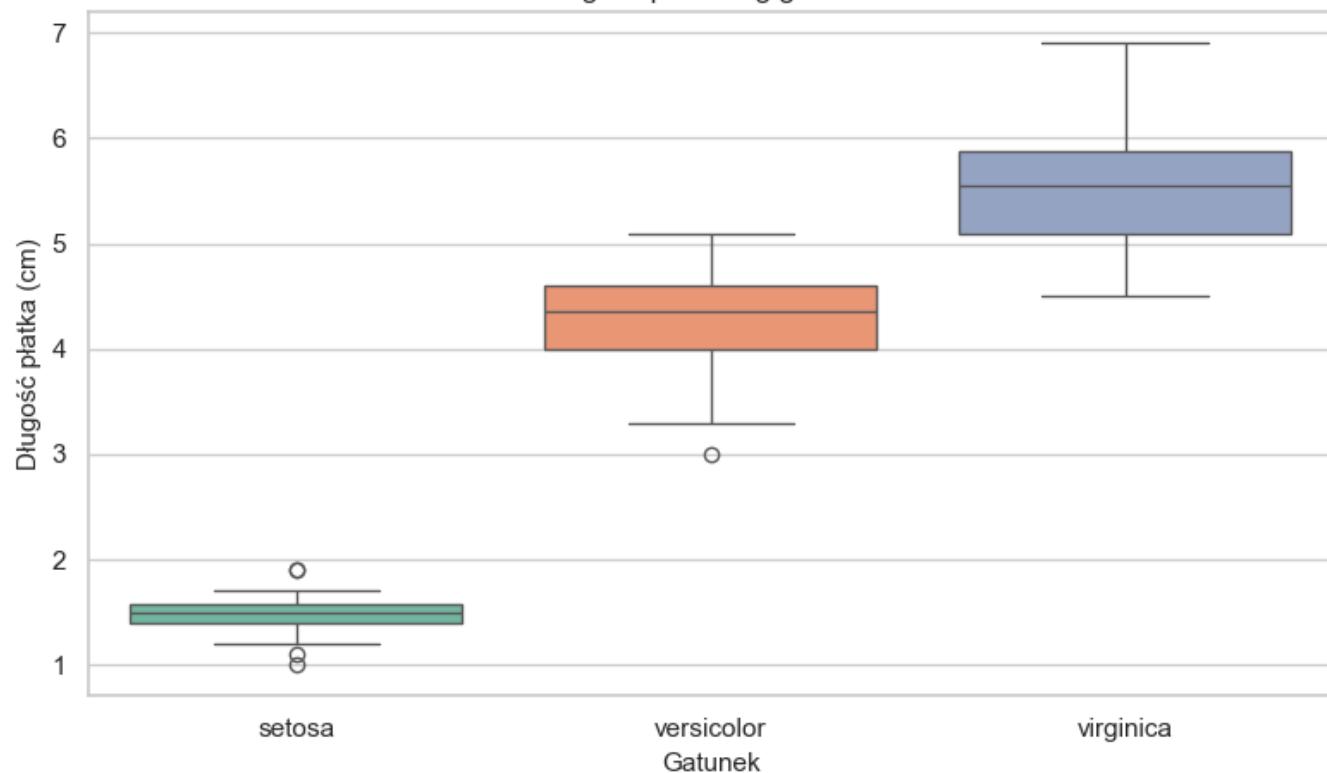
	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

Iris Setosa — najmniejsze płatki

Setosa wyróżnia się najmniejszymi płatkami — to najbardziej kompaktowy gatunek irysa, szczególnie pod względem **długości i szerokości płatka**.

Ta cecha powoduje, że *setosa* wyraźnie odróżnia się od pozostałych dwóch gatunków.

Długość płatka wg gatunku





Iris setosa

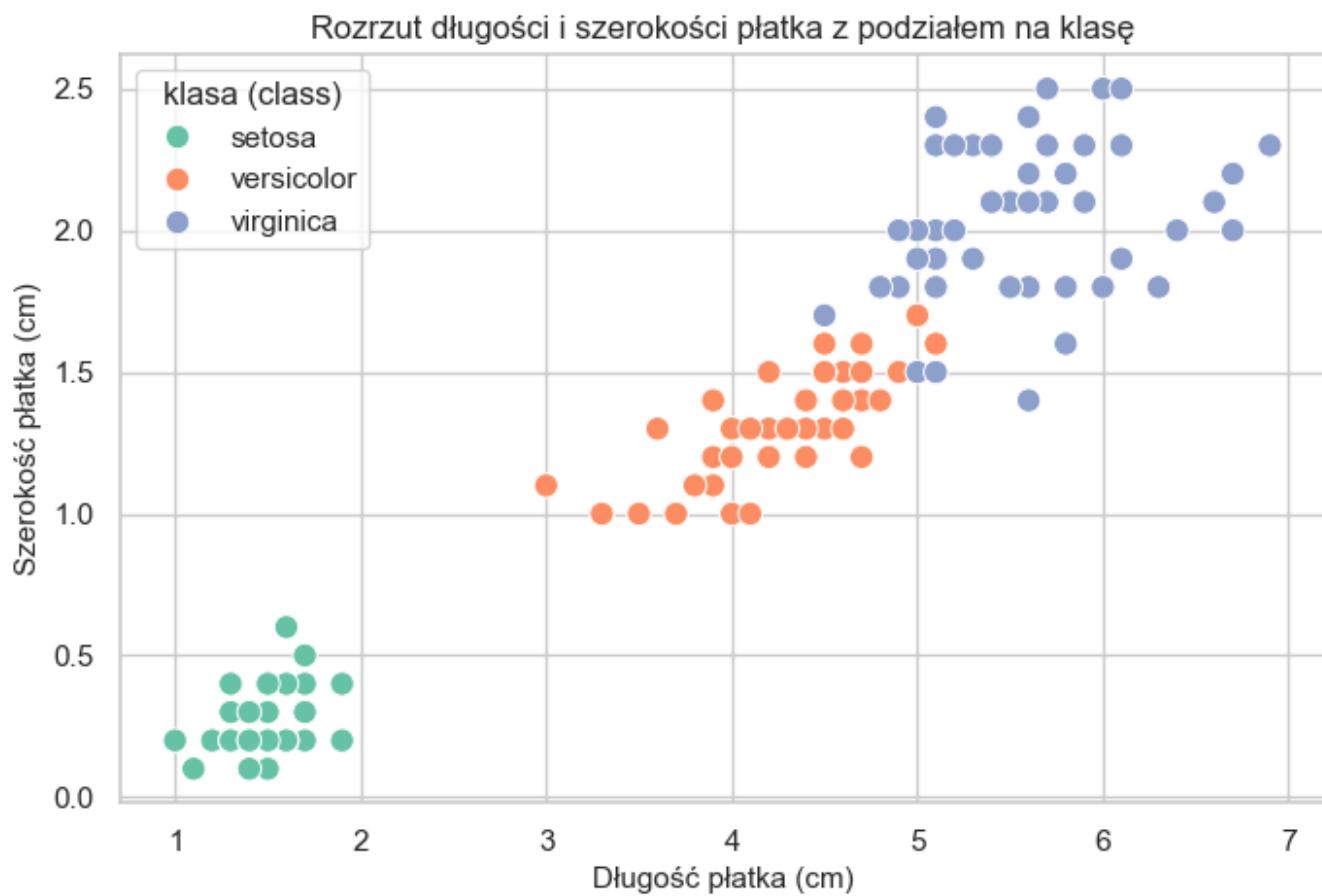
– Blue Flag Iris

Iris Versicolor — cechy pośrednie

Versicolor zajmuje pozycję pośrednią między setosą a virginicą pod względem rozmiaru płatków.

Widzimy, że:

- płatki są **większe niż u setosy**,
- ale **mniejsze niż u virginiki**,
- wartości układają się **wzdłuż przejścia** pomiędzy pozostałymi gatunkami.

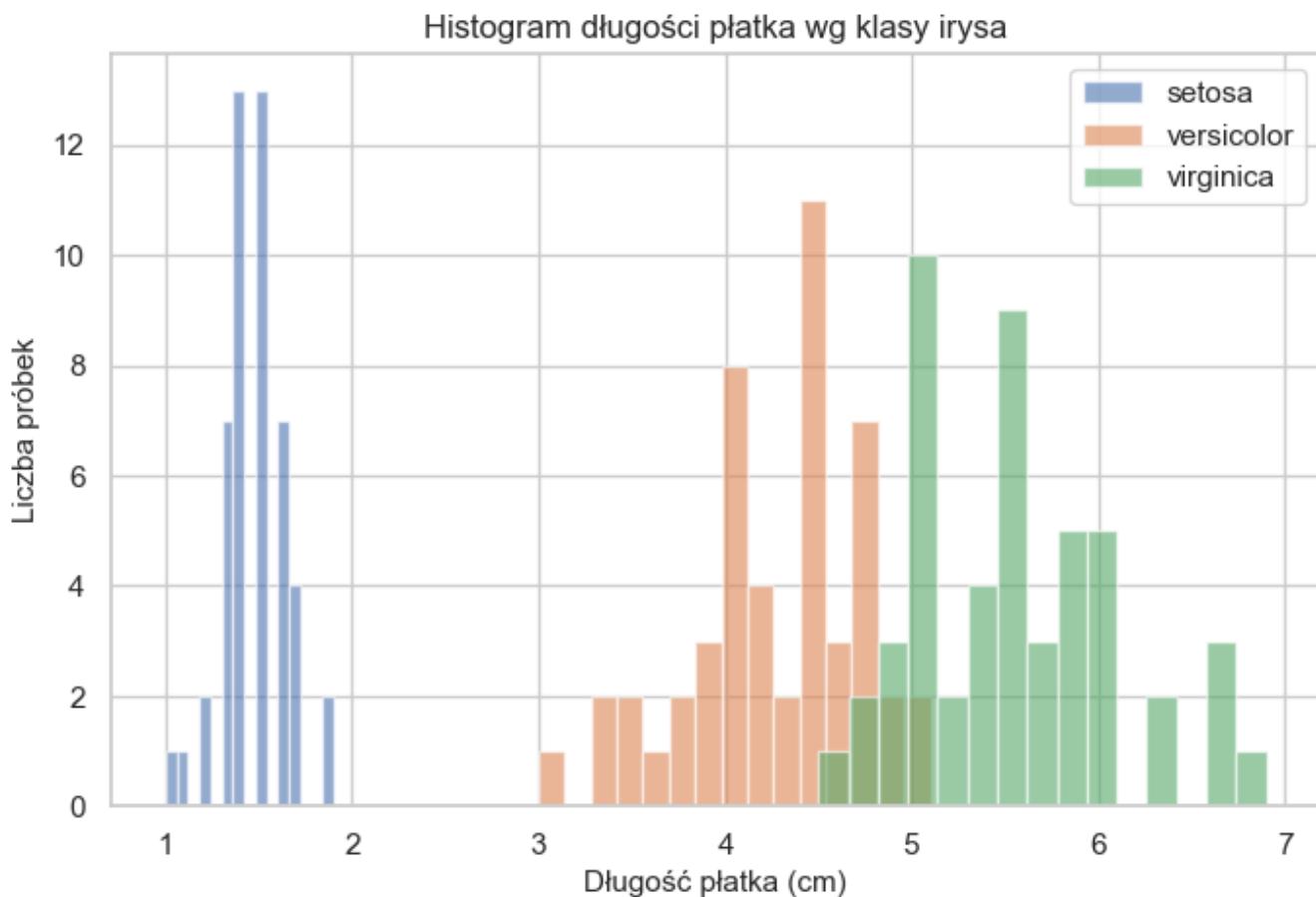




Iris versicolor – Harlequin Blueflag

Iris Virginica — największe płatki

Virginica ma **największe płatki** (długość i szerokość), co mocno odróżnia ją od pozostałych gatunków.





Iris virginica
Southern Blue Flag

Najważniejsze wnioski z analizy

Różnice między gatunkami

- **Setosa** wyróżnia się najmniejszymi płatkami (petal length, petal width)
- **Virginica** ma największe płatki — szczególnie szerokość płatka
- **Versicolor** zajmuje pozycję pośrednią między pozostałymi dwoma gatunkami

Te cechy pozwalają **wyraźnie oddzielić gatunki** w przestrzeni cech.

Śr. dł. płatka Śr. szer. płatka Śr. dł. kielicha

klasa (class)	średnia długość płatka	średnia szerokość płatka	średnia długość kielicha
setosa	1.46	0.25	5.01
versicolor	4.26	1.33	5.94
virginica	5.55	2.03	6.59

Śr. dł. płatka Śr. szer. płatka Śr. dł. kielicha

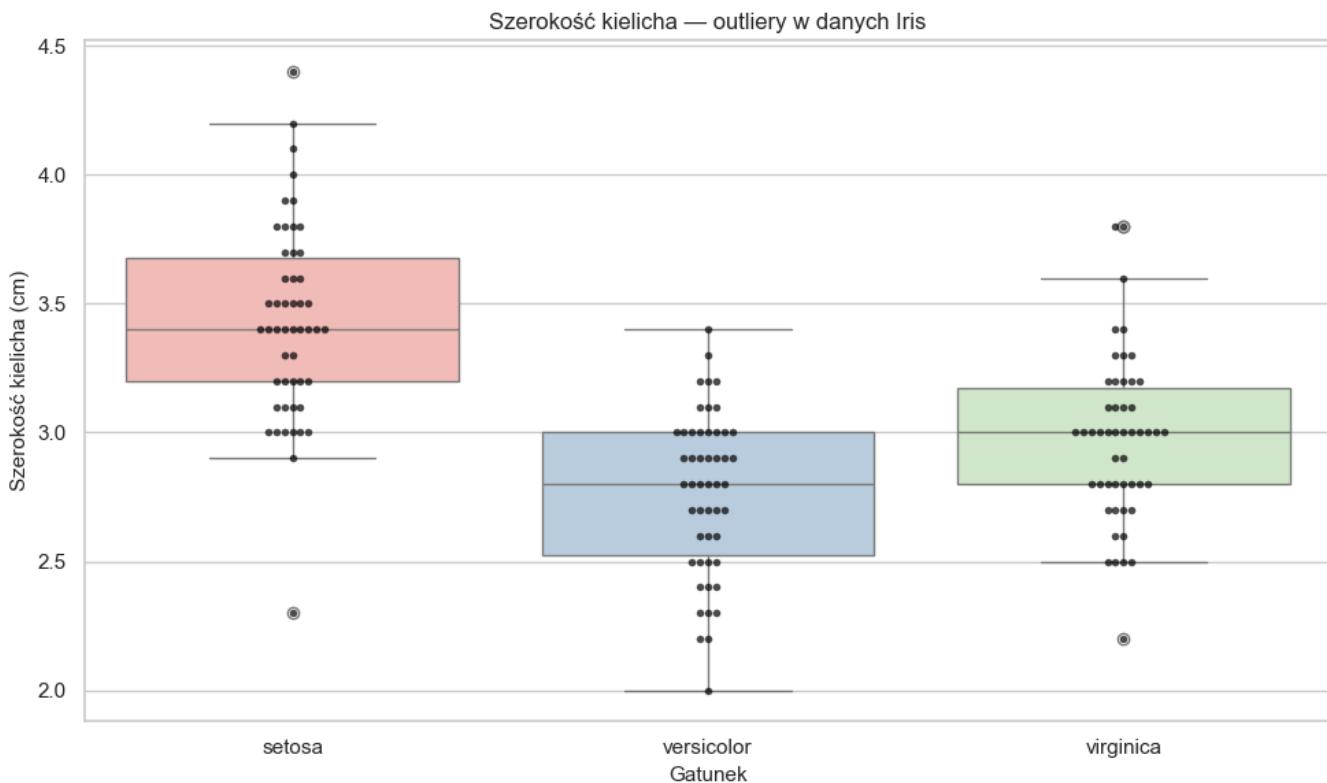
klasa (class)	średnia długość płatka	średnia szerokość płatka	średnia długość kielicha
setosa	1.460000	0.250000	5.010000
versicolor	4.260000	1.330000	5.940000
virginica	5.550000	2.030000	6.590000

Czy w danych są wartości odstające?

- W gatunku **setosa** pojawiają się pojedyncze kwiaty o nietypowej szerokości kielicha — jednak mieszczą się one w naturalnej zmienności.
- U **virginica** widać nieco większą różnorodność w szerokości płatków, co może wynikać z bardziej zróżnicowanej budowy tego gatunku.
- Pozostałe cechy są stabilne i układają się w bardzo czytelne wzorce między gatunkami.

Nietypowe wartości są **nieliczne** i wyglądają na **naturalne różnice biologiczne**, a nie błędy pomiarowe.

Dane są czyste i nie wymagają dodatkowego przetwarzania przed modelowaniem.



Co wynika z tej analizy?

Dane o irysach pokazują, że **wystarczy kilka prostych pomiarów**, aby z dużą pewnością odróżnić gatunki roślin.

Różnice są **wyraźne, naturalne biologicznie** i stabilne pomiędzy próbками.

To sprawia, że zbiór Iris:

- jest **idealnym przykładem** do nauki klasyfikacji
- świetnie sprawdza się w edukacji i projektach ML
- pozwala zrozumieć, jak cechy wpływają na rozpoznawanie obiektów

🎯 Podsumowanie

- płatki → najważniejsza cecha odróżniająca gatunki
- dane są czyste i gotowe do modelowania
- irisy = świetny przykład klasyfikacji w biologii i ML