

# Programowanie objaśniające ze Sweave

Przemysław Biecek

Wykres pudełkowy jest bardzo popularną metodą prezentacji zmienności pojedynczej zmiennej. Można go wyznaczać również dla kilku zmiennych (dzięki czemu możemy porównać rozkłady tych zmiennych) lub dla pojedynczej zmiennej w rozbiciu na grupy obserwacji.

Kształt tego wykresu jest bardzo charakterystyczny, przypominający pudełko z wąsami. Poszczególne elementy wykresu przedstawiają różne charakterystyki obserwowanej zmiennej. Środek pudełka przedstawia medianę, dolna i górna granica pudełka odpowiada kwartylom z próby (odpowiednio dolnemu i górnemu kwartylowi), kropki przedstawiają obserwacje odstające. Zakres zmienności (minimum i maksimum) danej zmiennej (po ominięciu wartości odstających) zaznaczony jest za pomocą wąsów wykresu.

```
> dane <- read.csv("http://www.biecek.pl/R/dane/daneSoc.csv", sep = ";")
> attach(dane)
> print(by(wiek, plec, summary))
```

INDICES: kobieta

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
23.00	32.00	47.00	46.38	57.00	75.00

INDICES: mezczyzna

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
22.00	30.00	43.00	41.97	53.00	74.00

```
> boxplot(wiek ~ plec, data = dane, col = "lightgrey")
```

