

## Ćwiczenie 8

Zakładamy, że człowiek ma nadciśnienie tętnicze, jeżeli jego ciśnienie skurczowe przekracza 140 mmHg, a ciśnienie rozkurczowe 100 mmHg.

Plik „cisnienie.txt” ma następującą strukturę kolumn:

- 1) ciśnienie skurczowe
- 2) ciśnienie rozkurczowe
- 3) płeć (1 – mężczyzna, 0 – kobieta)
- 4) nadużywanie papierosów (1 – tak, 0 – nie)
- 5) nadużywanie alkoholu (1 – tak, 0 – nie)
- 6) wiek.

### Cel ćwiczenia:

- 1) Zaproponować model regresji liniowej dla zmiennych zależnych „ciśnienie skurczowe” i „ciśnienie rozkurczowe”
- 2) Wybrać model z największą wartością współczynnika  $R^2$
- 3) Przeprowadzić testy istotności dla współczynników regresji (tylko dla modelu z największą wartością współczynnika  $R^2$ );
- 4) Sprawdzić, czy reszty mają rozkład normalny (tylko dla modelu z największą wartością współczynnika  $R^2$ ) – funkcja w **R** shapiro.test()
- 5) Obliczyć współczynnik korelacji dla zmiennych „ciśnienie skurczowe” i „ciśnienie rozkurczowe”
- 6) Utworzyć nową kolumnę danych o nazwie „nadciśnienie” w taki sposób że, człowiek mający problemy z ciśnieniem skurczowym lub rozkurczowym będzie przyjmował wartość 1, a człowiek nie mający nadciśnienia wartość 0
- 7) Utworzyć model regresji logistycznej dla zmiennej „nadciśnienie” i sprawdzić czy bardziej prawdopodobne jest że nadciśnienie ma kobieta nadużywająca alkoholu i papierosów czy mężczyzna nadużywający obydwu używek