

Ćwiczenie 2. (2pt) Obserwacje odstające

Następujące polecenia powinny być wykonane dla **dwóch wskazanych zmiennych** ze zbioru "cars_acceleration.csv".

Przydział zmiennych do grup studentów jest przedstawiony w pliku "Przydziały.docs".

Polecenia:

1. Zidentyfikuj obserwacje odstające dla wskazanych zmiennych używając:
 - a. histogramów
 - b. wykresów rozrzutu
2. Sprawdź, czy wartości wskazane w punkcie 1 są obserwacjami odstającymi używając:
 - a. metody standaryzacji (the Z-score method)
 - b. metody rozstępu międzykwartylowego (the IQR method)
3. Które z obserwacji odstających wskazanych metodami numerycznymi nie są widoczne na histogramach i wykresach rozrzutu? Czy na histogramach bądź wykresach rozrzutu są widoczne takie obserwacje odstające, które nie zostały wyznaczone przez metody numeryczne.
4. Prześledź, jak obserwacje odstające wpływają na wartości średniej i mediany. W tym celu:
 - a. Oblicz średnią i medianę dla danych z i bez obserwacji odstających.
 - b. Stwierdź, która z miar tendencji centralnej, średnia czy mediana, jest bardziej wrażliwa na obecność obserwacji odstających. Spróbuj wyjaśnić dlaczego.