

Ćwiczenie 1 (2pt). Obsługa brakujących danych

Następujące polecenia powinny być wykonane dla **wskazanej zmiennej** ze zbioru "data.csv".

Przydział zmiennych do grup studentów jest przedstawiony w pliku "Przydzialy.docx".

Polecenia:

1. Obliczyć wartość średnią i odchylenie standardowe po usunięciu brakujących (NA) wartości. Narysować histogram.
2. Zastąpić brakujące wartości wartościami średnimi obliczonymi w punkcie 1.
 - a. Narysować histogram.
 - b. Obliczyć średnią i odchylenie standardowe.
 - c. Porównać wyniki z wynikami z punktu 1.
3. Zastąpić brakujące wartości wartościami generowanymi losowo zgodnie z obserwowanym rozkładem zmiennej.
 - a. Narysować histogram.
 - b. Obliczyć średnią i odchylenie standardowe.
 - c. Porównać wyniki z wynikami z punktów 1 i 2.

Uwagi

Wyniki powinny być dostarczone w postaci **raportu**, w którym znajdą się odniesienia do wszystkich kolejnych punktów ćwiczenia. Wszystkie wyniki powinny zostać opisane i skomentowane (nie jest wymagane dostarczenie R kodu).

Fragment przykładowego raportu.

Zad. 1....

Zad. 2,3. ... Na rysunku??, przedstawiony jest histogram zmiennej ?? z brakującymi wartościami zastąpionymi przez ??. Porównując ten histogram z histogramem przedstawionym na rysunku?? można zauważyć, że Wartość średnia zmiennej?? dla tego przypadku jest (większa, mniejsza, prawie równa) średniej

Na podstawie wyników przeprowadzonego eksperymentu można stwierdzić, że (napisać, który sposób zastąpienia brakujących danych był lepszy).