

Zadanie 1

Print nie jest jedyną metodą pozwalającą na wyświetlenie danych. Spróbuj wyświetlić dane ze zmiennej `df` z wykorzystaniem `display` z biblioteki `pandas`. Chodzi o zmianę konfiguracji `display` i wyświetlenie danych. Proszę ustawić:

- Maksymalną liczbę wierszy
- Maksymalną liczbę kolumn
- Maksymalną liczbę wierszy do wyświetlenia (np. 1000)
- Wycentrować wartości kolumn
- Ustawić precyzję wyświetlania liczb dla całej tabeli do 3 miejsca po przecinku

Zadanie 2

Wypróbuj dwa inne sposoby na wypełnienie brakujących wartości. Wypełnij wszystkie brakujące wartości medianą i zapisz dataframe jako `df_3`. Wyświetl pod tabelą informację o medianie dla zmiennej `wiek` oraz `wzrost`. Wypełnij wszystkie brakujące wartości jako 0 i zapisz dataframe jako `df_4`.

Zadanie 3

Przełóż zdobytą wiedzę, żeby dokonać podstawowej eksploracji nowego zestawu danych. Wczytaj dane `DSP_5.csv` pamiętając o właściwym określeniu separatora. Sprawdź czy są brakujące dane - jeżeli są, to wypełnij je średnią. Przedstaw średnią, wariancję i rozstęp dla zmiennej `"hp"`. Przedstaw tabelę korelacji dla zmiennych. Czy są jakieś wartości, które szczególnie zwracają uwagę?

Zadanie 4

1. Wczytaj dane „`DSP_8.csv`”. Jest to zbiór danych Heart Attack Analysis & Prediction Dataset (aut. Rashik Rahman) dostępny w Kaggle. Zadbaj o odpowiednie formatowanie liczb dziesiętnych.
2. Napisz kod, który zwróci informację dotyczącą:
 - a) liczby kolumn (wraz z ich nazwami),
 - b) liczby wierszy (obserwacji),
 - c) ewentualnych braków danych,
 - d) średniego wieku oraz odchylenia standardowego w grupie kobiet oraz w grupie mężczyzn,
 - e) odsetka mężczyzn w zbiorze danych,
 - f) liczby kobiet w wieku od 45 do 50 lat,
 - g) korelacji pomiędzy zmiennymi, ale wyłącznie dla osób, dla których w zbiorze danych określono, że ich EKG w czasie spoczynku jest w normie (czyli w kolumnie `RestingECG` występuje obserwacja „Normal”).