

SPRAWOZDANIE

Zajęcia: Nauka o danych I

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium Nr 2 Data 18.10.2024 Temat: "Praktyczne Zastosowanie Podstawowych Funkcji Statystycznych w Analizie Danych" Wariant 1	Imię Nawisko: Mariusz Jagosz Informatyka II stopień, niestacjonarne, 1 semestr, gr.A
---	---

1. Polecenie: wariant 1 zadania

Oblicz podstawowe funkcje statystyczne używając zbiory danych z poprzedniego zajęcia

Dane znajdują się pod linkiem: Development Assistance for Health Database 1990-2020 <http://ghdx.healthdata.org/record/ihme-data/development-assistance-health-database-1990-2020>

2. Opis programu opracowanego (kody źródłowe, rzuty ekranu)

Kod znajduje się na repozytorium pod linkiem:

[unibb/Nauka o Danych I/Task2 at main · mariuszjagosz/unibb](#)

Kod został również przesłany na platformę elearningową

W celu wykonania kodu należy pobrać dane z powyższego linku i umieścić plik CSV w tym samym folderze, w którym znajduje się notebook Jupyter.

3. Wnioski

1. **Średnia i mediana:** Średnia wartości "dah_20" (2519,36) jest znacznie wyższa od mediany (44), co wskazuje na skośny rozkład z obecnością ekstremalnych wartości.

2. **Odchylenie standardowe i wariancja:** Wysoka zmienność (odchylenie standardowe: 36693,26) sugeruje dużą różnorodność w danych, co może wynikać z różnych poziomów pomocy dla krajów.
3. **Korelacja:** Silna dodatnia korelacja (0,75) między "other_dah_20" a "dah_20" wskazuje na współzależność między tymi dwoma rodzajami pomocy.
4. **Kowariancja:** Dodatnia kowariancja (213078437,00) potwierdza, że wyższe wartości w jednej kategorii wiążą się z wyższymi wartościami w drugiej.

Wnioski te sugerują, że pomoc w tych dwóch kategoriach jest przydzielana w sposób skoordynowany i różni się między krajami.