

# Raport domowy nr 1

1 grudnia 2023

**Temat raportu:** Wykorzystanie poznanych metod służących do analizy zależności liniowej dla wybranych danych rzeczywistych.

**Termin oddania:** 22.12.2023, godzina 23:59.

**Raport wykonujemy w dwuosobowych grupach.** Pracujecie Państwo na dowolnie wybranych danych. Należy wybrać dane, które są zależne liniowo (lub dla których otrzymujemy zależność liniową po transformacji danych).

**Raport powinien zawierać następujące rozdziały:**

1. Opis danych.
2. Analiza jednowymiarowa zmiennej zależnej oraz zmiennej niezależnej:
  - wizualizacja danych (na przykład boxplot, histogram, gęstość, dystrybucja),
  - podstawowe miary położenia, rozproszenia, skośności, spłaszczenia,
  - interpretacja wyników.
3. Analiza zależności liniowej pomiędzy zmienną zależną a zmienną niezależną:
  - prezentacja danych (wykres rozproszenia) i określenie zależności,
  - estymacja współczynników w klasycznym modelu regresji (estymacja punktowa oraz estymacja przedziałowa),
  - ocena poziomu zależności (współczynnik korelacji Pearsona, SST, SSE, SSR),
  - predykcja oraz przedziały ufności dla pewnej części danych,
  - interpretacja wyników.
4. Analiza residuów.
  - sprawdzenie założeń: średnia równa zero, stała wariancja, niezależność residuów, rozkład normalny,
  - interpretacja wyników.
5. Podsumowanie i wnioski.

### **Dodatkowe uwagi:**

- Proszę zwrócić uwagę na odpowiedni styl i właściwą formę raportu: podział na rozdziały/sekcje, a także na odpowiednie opracowanie i przedstawienie tabeli oraz wykresów. Do każdej tabeli oraz każdego wykresu powinno znaleźć się odwołanie w tekście. Należy je również odpowiednio sformatować (zadbać o to, by wykresy i tabele były czytelne oraz zawierały podpisy).
- Raport powinien zawierać opis danych (informację czego dane dotyczą oraz jak długa jest próba). W pracy należy wskazać skąd pochodzą dane (jeśli jest to możliwe, dobrze jest umieścić w pracy link do strony, z której można pobrać dane).
- Jeżeli korzystają Państwo z jakiś statystyk, wzorów czy twierdzeń to proszę je zamieścić w pracy (oraz odpowiednio zacytować). Np. w przypadku testów statystycznych należy opisać na czym polega test, wskazać hipotezę zerową i hipotezę alternatywną oraz zamieścić wzór statystyki testowej.
- Jeśli jest taka potrzeba to przed rozpoczęciem analiz dane można wygładzić, lub dokonać ich transformacji.
- Można spróbować dodatkowo usunąć wartości odstające i sprawdzić jak to wpłynęło na jakość dopasowania (względem dopasowania dla danych z obecnością wartości odstających). W takim przypadku należy opisać działanie metody do usuwania wartości odstających oraz uzasadnić jej zastosowanie dla omawianych danych.
- W przypadku badania normalności rozkładu poza testami wizualnymi należy wykonać również testy statystyczne.
- W pracy oprócz rezultatów powinny zostać umieszczone również odpowiednie wnioski i interpretacja otrzymanych wyników – dotyczy to każdej z trzech części raportu.