

# Wnioskowanie i klasyfikacja w analizie danych

## Projekt 2. Jednoczynnikowa analiza wariancji (ANOVA).

W załączonym pliku dane13.csv zapisano 40 wartości, które zostały pogrupowane w trzech kategoriach. Każdy student losuje 15 liczb z rozkładu normalnego  $\mathcal{N}(10.2, (3.8)^2)$  i dopisuje otrzymane wartości do zbioru w dane13.csv jako czwartą klasę (kategorię). Na potrzeby zajęć projektowych uznajemy, że otrzymany w ten sposób zbiór danych jest  $n = 55$ -elementową populacją próby podzieloną ze względu na pewien czynnik na 4 grupy (kategorie).

Pytamy, czy średnia wartość obserwowanej cechy ilościowej różni się w różnych grupach. W celu zbadania wpływu zmiennej niezależnej (czynnika) na zmienną zależną, proszę przeprowadzić analizę wariancji (jednoczynnikową ANOVA).

W raporcie należy wykonać podane poniżej polecenia. Proszę o zachowanie kolejności prezentowanych wyników analizy, co ułatwi ocenę wykonanego projektu.

- Przedstaw statystyki opisowe dla każdej z czterech grup (medianę, kwartyle pierwszy i trzeci, wartość minimalną, wartość maksymalną, średnią, wariancję). Wykonaj wykres pudełkowy dla każdej z grup.
- Przedstaw tabelę jednoczynnikowej analizy wariancji.
- Wykonaj testy post hoc (o ile to konieczne).
- Sprawdź założenia ANOVA i zwróć uwagę na konsekwencje ich naruszania.

Proszę pamiętać o prezentowaniu wniosków z wykonanych analiz.

Raport powinien być zapisany w profesjonalnym edytorze typu LaTeX.

Ważne. Wraz z raportem z wykonanego projektu proszę załączyć plik z danymi, dla których przeprowadzona została analiza. Wartości należące do klasy 4 (wraz z numerem klasy) powinny być dopisane poniżej, tj. zaczynając od wiersza nr 42 i z zachowaniem odpowiedniego formatu. W przypadku braku pliku lub w przypadku nieodpowiedniego formatu pliku projekt nie zostanie oceniony.