

**Karta przedmiotu oferowanego w Szkole Doktorskiej nr 3**  
**– semestr letni 2021/2022**

TYTUŁ
Analiza i wizualizacja danych w języku R
JEDNOSTKA PROWADZĄCA
Szkoła Doktorska nr 3
DYSCYPLINA NAUKOWA
Nauki fizyczne
JEDNOSTKA REALIZUJĄCA
105000 - Wydział Fizyki
OPIS PRZEDMIOTU
<p>Język R jest otwartym, darmowym i szybko rozwijającym się środowiskiem przeznaczonym przede wszystkim do analizy i wizualizacji danych. Dzięki bardzo łatwej składni i wielkiej liczbie bibliotek jest wykorzystywany zarówno w nauce jak i w biznesie przez zespoły badawczo-rozwojowe na całym świecie. Celem przedmiotu jest zapoznanie słuchaczy z podstawami języka R oraz nauczenie ich metod analizy i wizualizacji danych w tym języku.</p> <p>Treści kształcenia</p> <ol style="list-style-type: none"><li>1. Wprowadzenie do języka R.<ol style="list-style-type: none"><li>a) Zmienne, instrukcje sterujące, zapis i odczyt danych do pliku.</li><li>b) Operacje na wektorach i macierzach.</li><li>c) Generacja liczb losowych z zadanego rozkładu.</li><li>d) Generowanie raportów w języku Markdown.</li><li>e) Połączenie i praca z bazą danych MySQL.</li></ol></li><li>2. Wizualizacja w języku R.<ol style="list-style-type: none"><li>a) Wykresy XY i skrzynkowe.</li><li>b) Histogramy, binowanie.</li><li>c) Biblioteka ggplot2.</li><li>d) Generacja plików w formatach PNG, JPG i EPS.</li></ol></li><li>3. Wstępne przetwarzanie danych, podstawowe testy i modele statystyczne.<ol style="list-style-type: none"><li>a) Transformacja danych.</li><li>b) Testowanie zgodności hipotezy metodami parametrycznymi i nieparametrycznymi.</li><li>c) Dopasowywanie rozkładów teoretycznych do danych eksperymentalnych.</li><li>d) Test chi2.</li><li>e) Jedno i wielowymiarowa analiza wariancji.</li></ol></li></ol>

f) Regresje liniowa, logistyczna i Poissona.  
4. Eksploracja danych (uczenie maszynowe).

- a) Skuteczność klasyfikacji.
- b) Walidacja krzyżowa.
- c) Analiza dyskryminacyjna.
- d) Analiza skupień.
- e) Skalowanie wielowymiarowe.
- f) Analiza składowych głównych.
- g) Analiza przeżycia.

#### Metody kształcenia

Zajęcia mają formę warsztatów (laboratoriów) z użyciem komputerów. Podczas pierwszej części zajęć prowadzący wprowadza nowe treści i ilustruje je przykładami, natomiast w trakcie drugiej części słuchacze samodzielnie rozwiązują zadanie punktowane. Zadania punktowane mogą być dokańczane również po zajęciach.

#### Metody i kryteria oceniania

Zadania punktowane: 10 x 5 pkt = 50 pkt

Kolokwium końcowe z zadań praktycznych: 20 pkt

Razem do uzyskania: 70 pkt

Oceny: 35-41 dst, 42-48 dst+, 49-55 db, 56-62 db+, 63-70 bdb

### LITERATURA

- [1] M. Gogolewski, Programowanie w języku R, Wydawnictwo Naukowe PWN, Warszawa 2014
- [2] P. Biecek, Przewodnik po pakiecie R, Oficyna Wydawnicza Gis, Wrocław 2011.
- [3]. T. Górecki, Podstawy statystyki z przykładami w R, Wydawnictwo BTC, Legionowo 2011.
- [4] M. Walesiak, E. Gatnar , Statystyczna analiza danych z wykorzystaniem pakietu R, Wydawnictwo Naukowe PWN, Warszawa 2013.
- [5] A. Petrie, C. Sabin, Statystyka medyczna w zarysie, Wydawnictwo Lekarskie PZWL, Warszawa 2006.
- [6] K. Seefeld, E. Linder, Statistics Using R with Biological Examples, [https://cran.r-project.org/doc/contrib/Seefeld\\_StatsRBio.pdf](https://cran.r-project.org/doc/contrib/Seefeld_StatsRBio.pdf)
- [7] A. Coghlan, A Little Book of R For Biomedical Statistics, <http://a-little-book-of-r-for-biomedical-statistics.readthedocs.io/en/latest/>
- [8] A. Coghlan, A Little Book of R for Multivariate Analysis, <http://little-book-of-r-for-multivariate-analysis.readthedocs.io/en/latest/>

### EFEKTY UCZENIA

Student, który zaliczył przedmiot:

- Zna i rozumie podstawowe metody analizy i wizualizacji danych (P8S\_WG).
- Potrafi wnioskować na podstawie wyników uzyskanych podczas analizy i wizualizacji danych (P8S\_UW).
- Potrafi wygenerować raport w języku Markdown prezentujący uzyskane wyniki i wnioski (P8S\_UK).
- Jest gotów do samodzielnego prowadzenia badań naukowych obejmujących analizę i wizualizację danych (P8S\_KR)

METODY I KRYTERIA OCENIANIA ORAZ FORMA ZALICZENIA ZAJĘĆ		
Zadania punktowane: 10 x 5 pkt = 50 pkt Kolokwium końcowe z zadań praktycznych: 20 pkt Razem do uzyskania: 70 pkt Oceny: 35-41 dst, 42-48 dst+, 49-55 db, 56-62 db+, 63-70 bdb		
JĘZYK WYKŁADOWY PRZEDMIOTU		PUNKTY ECTS
polski		2
FORMA PROWADZONYCH ZAJĘĆ	WYMIAR GODZIN	PROWADZĄCY
Laboratorium (LAB)	30	Robert Paluch, dr inż.