

## Wprowadzenie do R, Laboratorium, kurs zaawansowany

### Karta opisu przedmiotu

#### Informacje podstawowe

<p><b>Kierunek studiów</b> Socjologia</p> <p><b>Ścieżka</b> Analiza danych i badania społeczne</p> <p><b>Jednostka organizacyjna</b> Wydział Filozoficzny</p> <p><b>Poziom kształcenia</b> drugiego stopnia</p> <p><b>Forma studiów</b> studia stacjonarne</p> <p><b>Profil studiów</b> ogólnoakademicki</p> <p><b>Obligatoryjność</b> fakultatywny</p>	<p><b>Cykl kształcenia</b> 2023/24</p> <p><b>Kod przedmiotu</b></p> <p><b>Języki wykładowe</b> Polski</p> <p><b>Dyscypliny</b> Nauki socjologiczne</p> <p><b>Klasyfikacja ISCED</b> 0314 Socjologia i kulturoznawstwo</p> <p><b>Kod USOS</b></p>
<b>Koordinator przedmiotu</b>	Maria Varlamova
<b>Prowadzący zajęcia</b>	Maria Varlamova

<b>Okresy</b> semestr 2, semestr 4	<p><b>Forma weryfikacji uzyskanych efektów uczenia się</b> zaliczenie na ocenę</p> <p><b>Sposób realizacji i godziny zajęć</b> ćwiczenia laboratoryjne: 15</p>	<b>Liczba punktów ECTS</b> 2.0
---------------------------------------	--	-----------------------------------

#### Cele kształcenia dla przedmiotu

C1	Nabycie dogłębnej wiedzy na temat zaawansowanych funkcji programowania w R, w tym manipulacji danymi, wizualizacji i modelowania statystycznego.
C2	Rozwinięcie umiejętności stosowania zaawansowanych metod statystycznych i analizy danych w kontekście badań socjologicznych.

C3	Nauczenie się analizowania i interpretowania złożonych zestawów danych socjologicznych, wykorzystując różnorodne techniki statystyczne i algorytmy uczenia maszynowego.
C4	Zdobycie praktycznych umiejętności w zakresie stosowania technik uczenia maszynowego do rozwiązywania rzeczywistych problemów socjologicznych.
C5	Wzmocnienie umiejętności krytycznego myślenia i analizy danych, umożliwiające studentom ocenę i interpretację wyników badań w kontekście teoretycznym i praktycznym.

### Efekty uczenia się dla przedmiotu

Kod	Efekty w zakresie	Kierunkowe efekty uczenia się	Metody weryfikacji
<b>Wiedzy – Student zna i rozumie:</b>			
W1	Student pogłębia wiedzę na temat zaawansowanych funkcji środowiska R, w tym programowania, analizy danych i technik modelowania statystycznego.	SOC_K2_W06	projekt
W2	Student zdobywa specjalistyczną wiedzę z zakresu zastosowania metod statystycznych i technik uczenia maszynowego w analizie danych socjologicznych.	SOC_K2_W08	projekt
<b>Umiejętności – Student potrafi:</b>			
U1	Student umie zaawansowanie przetwarzać i analizować dane socjologiczne w R, wykorzystując nowoczesne pakiety i techniki programistyczne.	SOC_K2_U04	projekt
U2	Student jest w stanie wykorzystać zaawansowane metody statystyczne do identyfikacji, analizy i interpretacji złożonych wzorców i trendów w danych socjologicznych.	SOC_K2_U03	projekt
U3	Student potrafi samodzielnie stosować techniki uczenia maszynowego i modelowania statystycznego do przewidywania i modelowania zjawisk społecznych.	SOC_K2_U02	projekt
<b>Kompetencji społecznych – Student jest gotów do:</b>			
K1	Student rozwija umiejętność krytycznego myślenia i etycznego podejścia do analizy danych, z uwzględnieniem aspektów społecznych i kulturowych.	SOC_K2_K01	projekt
K2	K2: Student jest przygotowany do efektywnego komunikowania wyników analiz i badań w środowisku naukowym i zawodowym, oraz do współpracy w zespołach badawczych.	SOC_K2_K02	projekt

## Bilans punktów ECTS

Forma aktywności studenta	Średnia liczba godzin* przeznaczonych na zrealizowane rodzaje zajęć	
ćwiczenia laboratoryjne	15	
przygotowywanie projektów	20	
rozwiązywanie zadań	15	
Konsultacje	10	
<b>Łączny nakład pracy studenta</b>	<b>Liczba godzin</b> 45	<b>ECTS</b> 1.5
<b>Liczba godzin kontaktowych</b>	<b>Liczba godzin</b> 15	<b>ECTS</b> 0.5

\* godzina (lekcyjna) oznacza 45 minut

## Treści programowe

Lp.	Treści programowe	Efekty uczenia się dla przedmiotu
1.	Zaawansowane techniki czyszczenia danych (3 godziny) Dogłębne omówienie zaawansowanych metod czyszczenia danych, radzenie sobie z wartościami odstającymi, niespójnościami i skomplikowanymi przypadkami brakujących danych.	W1, U2, U3, K1
2	Analiza Głównych Składowych (PCA) i Klasteryzacja (3 godziny) Eksploracja PCA dla redukcji wymiarowości i różne techniki klasteryzacji (takie jak K-średnich, hierarchiczna) w analizie danych socjologicznych.	W2, U1, K1
3	Modele Efektów Mieszanych (3 godziny) Zrozumienie i stosowanie modeli efektów mieszanych dla danych o strukturze hierarchicznej lub grupowej.	W2, U1, U3, K1
4	Nadzorowane Uczenie Maszynowe: LDA, QDA i Drzewa Decyzyjne (3 godziny) Omówienie Analizy Dyskryminacyjnej Liniowej (LDA), Analizy Dyskryminacyjnej Kwadratowej (QDA) i Drzew Decyzyjnych do problemów klasyfikacji w socjologii.	W2, U1, U2, K1
5	Nienadzorowane Uczenie Maszynowe: Regresja Liniowa, Logistyczna i Wielowymiarowa; Las Losowy (3 godziny) Zajęcia na temat technik regresji i Lasu Losowego jako narzędzia uczenia nienadzorowanego, skoncentrowane na zastosowaniach socjologicznych.	W2, U1, U3, K1

## Informacje rozszerzone

### Metody nauczania:

seminarium, rozwiązywanie zadań, konsultacje

Rodzaj zajęć	Formy zaliczenia	Warunki zaliczenia przedmiotu
Seminarium i praktyczne ćwiczenia	Projekt badawczy: Studenci będą musieli wykonać projekt badawczy, który będzie polegać na zastosowaniu zaawansowanych technik R do analizy danych socjologicznych.	Student zaliczy projekt badawczy, jeśli przeprowadzi dokładną analizę danych, przygotuje naukowy raport i przedstawi prezentację.

---

## Wymagania wstępne i dodatkowe

Posiadanie podstawowej wiedzy z zakresu statystyki oraz znajomość języka programowania R.

### Literatura

#### Obowiązkowa

1. Teetor, P., & Loukides, M. K. (2011). R Cookbook : Proven Recipes for Data Analysis, Statistics, and Graphics (Vol. 1st ed). Sebastopol, CA: O'Reilly Media. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=edsebk&AN=414829>

#### Dodatkowa

1. Wickham, Hadley, and Garrett Golemund. 2016. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. " O'Reilly Media, Inc."
2. Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning (<http://faculty.marshall.usc.edu/gareth-james/>)
3. Colin Gillespie, Robin Lovelace, Efficient R programming (<https://bookdown.org/csgillespie/efficientR/>)