

WYDZIAŁ MATEMATYKI
KARTA PRZEDMIOTU

Nazwa przedmiotu w języku polskim: **Pakiety statystyczne**
 Nazwa przedmiotu w języku angielskim: **Statistical Packages**
 Kierunek studiów: **Matematyka stosowana**
 Specjalność:
 Stopień studiów i forma: **I stopień, stacjonarna**
 Rodzaj przedmiotu: **wybieralny**
 Kod przedmiotu:
 Grupa kursów: **TAK**

	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium
Liczba godzin zajęć zorganizowanych w Uczelni (ZZU)	30		30		
Liczba godzin całkowitego nakładu pracy studenta (CNPS)	175				
Forma zaliczenia	Egzamin/ zaliczenie na ocenę*	Egzamin/ zaliczenie na ocenę*	Egzamin/ zaliczenie na ocenę*	Egzamin/ zaliczenie na ocenę*	Egzamin/ zaliczenie na ocenę*
Dla grupy kursów zaznaczyć kurs końcowy (X)	X				
Liczba punktów ECTS	7				
w tym liczba punktów odpowiadająca zajęciom o charakterze praktycznym (P)	4				
w tym liczba punktów ECTS odpowiadająca zajęciom wymagającym bezpośredniego udziału nauczycieli lub innych osób prowadzących zajęcia (BU)	2,5				

*niepotrzebne skreślić

WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I KOMPETENCJI SPOŁECZNYCH

1. Student zna podstawy rachunku prawdopodobieństwa – *Rachunek prawdopodobieństwa.*

CELE PRZEDMIOTU

- C1 Poznanie składni oraz możliwości pakietu statystycznego R.
 C2 Zwrócenie uwagi na prawidłowe korzystanie z pakietu, sprawdzanie założeń metod oraz unikanie błędnych interpretacji wyników statystycznych.

PRZEDMIOTOWE EFEKTY UCZENIA SIĘ

Z zakresu wiedzy:

PEU_W01 Student zna powiązania pomiędzy terminami statystycznymi i statystyką matematyczną a praktyką stosowania pakietu statystycznego.

PEU_W02 Student odróżnia korelację od przyczynowości oraz zna znaczenie wybranej metody próbkowania oraz randomizacji dla rodzaju wniosków, jakie może wyciągnąć z analizy.

PEU_W03 Student zna wady i zalety systemu publikacji wyników opartych na analizach statystycznych, zna pojęcia obciążenia publikacyjnego, prerejestracji i replikacji badań oraz metaanalizy.

PEU_W04 Student zna założenia ważniejszych metod statystycznych oraz potrafi dobrać metodę do sytuacji.

Z zakresu umiejętności:

PEU_U01 Student potrafi sprawdzić założenia wybranych metod w pakiecie R oraz prawidłowo interpretować wyniki analiz statystycznych.

PEU_U02 Student potrafi wykonać analizę danych w pakiecie statystycznym R.

PEU_U03 Student potrafi zwizualizować dane w pakiecie statystycznym R.

PEU_U04 Student potrafi przygotować raport ze statystycznej analizy danych, cechujący się odpowiednią strukturą.

Z zakresu kompetencji społecznych:

PEU_K01 Student jest przygotowany do zdobywania nowych kompetencji i samodzielnej analizy danych w poszukiwaniu odpowiedzi na interesującego go pytania.

PEU_K02 Student jest przygotowany do współpracy z przedstawicielami innych dyscyplin naukowych.

PEU_K03 Student jest bardziej świadomym konsumentem informacji statystycznej, potrafi wykryć niektóre manipulacje, błędy i nadinterpretacje wyników statystycznych.

PEU_K04 Student komunikuje wyniki badań oraz analiz w sposób zrozumiały, formalny i ze zwróceniem szczególnej uwagi na poprawność wyciągniętych wniosków.

TREŚCI PROGRAMOWE

Forma zajęć - wykład		Liczba godzin
Wy1	Różne podejścia do statystyki; Statystyka w szerszym sensie a statystyka matematyczna i uczenie maszynowe; Praktyczne i techniczne aspekty pracy z pakietami statystycznymi.	2
Wy2	Przegląd cech statystycznych.	2
Wy3	Metody estymacji i metody oceny estymatorów oraz klasyfikatorów.	2
Wy4	Estymatory niektórych spośród omawianych cech.	2
Wy5	Podstawy grafiki statystycznej.	2
Wy6	Wstęp do testowania hipotez: podejście Fishera a Neymana-Pearsona, podstawowe pojęcia, błędy w interpretacji p-wartości, obciążenie publikacyjne, badania replikacyjne, prerejestracja badań, pojęcie metaanalizy.	1
Wy7	Testowanie hipotez: konstrukcja testów dla parametrów rozkładu normalnego (testy z, t, chi-kwadrat oraz F).	3

Wy8	Testowanie hipotez: konstrukcja testu parametru rozkładu wykładniczego, testy sumy rang i rang znakowanych Wilcoxon, testowanie normalności.	2
Wy9	Testowanie hipotez: testy dla korelacji, testy dla danych kategoriycznych (test Fishera, test G, testy chi-kwadrat Pearsona i McNemara).	2
Wy10	Regresja liniowa w ujęciu teorii najmniejszych kwadratów.	2
Wy11	Regresja liniowa w ujęciu modelu normalnego.	2
Wy12	Model liniowy, ważne przypadki szczególne: analiza wariancji, analiza kowariancji.	2
Wy13	Zarys podejścia bayesowskiego.	2
Wy14	Przedstawienie wybranego modelu statystycznego.	2
Wy15	Kolokwium zaliczeniowe.	2
	Suma godzin	30

Forma zajęć - laboratorium		Liczba godzin
La1	Omówienie środowiska R, jego wad i zalet w porównaniu z innymi pakietami statystycznymi oraz językami programowania; Korzystanie z programu RStudio; Przegląd interfejsu użytkownika; Sprawdzanie dokumentacji; Podstawowe obliczenia matematyczne.	2
La2	Podstawowe typy danych w R; Tworzenie, indeksowanie i modyfikowanie wektorów; Macierze; Instrukcje sterujące.	2
La3	Funkcje i podstawy programowania funkcyjnego; Ramki danych i operacje na nich; Zliczanie elementów do tabel.	2
La4	Wczytywanie danych; Statystyka opisowa w R; Identyfikowanie potencjalnych problemów podczas wczytywania zbiorów danych.	2
La5	Wizualizacja danych w pakiecie R.	2
La6	Ekosystem pakietów tidyverse.	2
La7	Testowanie hipotez w pakiecie R; Analiza mocy testów.	2
La8	Listy w pakiecie R; Klasy i metody; Implementacja własnego testu integrującego się ze standardowymi funkcjami pakietu R.	2
La9	Badanie własności testów za pomocą metod Monte Carlo; Własności asymptotyczne oszacowań wyliczanych metodą Monte Carlo; Dobór liczby replikacji.	2
La10	Wybór metody na podstawie symulacji; Porównanie przedziałów ufności dla parametru prawdopodobieństwa.	2
La11	Budowa modelu regresji liniowej dla wybranego zbioru danych; Weryfikacja założeń modelu; Analiza obserwacji odstających, o wysokiej dźwigni oraz wpływowych.	2
La12	Interpretacja wyników dopasowania modelu; Testowanie hipotez o parametrach oraz porównywanie modeli.	2
La13	Analiza wariancji; Ćwiczenia ilustrujące problem porównań wielokrotnych.	2
La14	Metody doboru zmiennych do modelu; Inżynieria cech.	2
La15	Ilustracja przykładowego modelu bayesowskiego za pomocą pakietu rstan.	2
	Suma godzin	30

STOSOWANE NARZĘDZIA DYDAKTYCZNE
N1. Wykład multimedialny z elementami tradycyjnego. N2. Laboratorium komputerowe. N3. Praca własna studenta.

OCENA OSIĄGNIĘCIA PRZEDMIOTOWYCH EFEKTÓW UCZENIA SIĘ

Oceny (F – formująca (w trakcie semestru), P – podsumowująca (na koniec semestru))	Numer efektu uczenia się	Sposób oceny osiągnięcia efektu uczenia się
F1	PEU_W01 PEU_W04	Prezentacja indywidualna podczas laboratoriów.

	PEU_K02 PEU_K04	
F2	PEU_W01 PEU_W04 PEU_U01 PEU_U02 PEU_U03 PEU_U04 PEU_K01 PEU_K02 PEU_K04	Sprawozdania wykonywane indywidualnie przez studentów.
F3	PEU_W02 PEU_W03 PEU_W04 PEU_U01 PEU_K01 PEU_K03	Kolokwium podczas wykładu.
$P = 5/7 * (0.3 * F1 + 0.7 * F2) + 2/7 * F3$		

LITERATURA PODSTAWOWA I UZUPEŁNIAJĄCA

LITERATURA PODSTAWOWA:

- [1] A. Reinhart, *Statistics Done Wrong*, No Starch Press 2015, wydanie 1.
- [2] H. Wickham, M. Çetinkaya-Rundel, G. Grolemund, *R for Data Science*, O'Reilly 2023, wydanie 2.
- [3] H. Wickham, *Advanced R*, CRC Press 2019, wydanie 2.
- [4] E. Jones, S. Harden, M. J. Crawley, *The R Book*, Wiley 2022, wydanie 3.

LITERATURA UZUPEŁNIAJĄCA:

- [1] D. S. Moore, G. P. McCabe, B. Craig, *Introduction to the Practice of Statistics*, W. H. Freeman 2021, wydanie 10.
- [2] F. Ramsey, D. Schafer, *The Statistical Sleuth: A Course in Methods of Data Analysis*, Cengage Learning 2012, wydanie 3.
- [3] E. R. Tufte, *The Visual Display of Quantitative Information*, Graphics Press 2001, wydanie 2.
- [4] S. Goodman, *A dirty dozen: twelve p-value misconceptions*, Semin Hematol 2008 Jul;45(3):135-40. doi: 10.1053/j.seminhematol.2008.04.003.
- [5] E. Gentle, W. K. Härdle, Y. Mori, *Handbook of Computational Statistics*, Springer 2012, wydanie 2.
- [6] J. Pearl, *Book of Why*, Basic Books 2018, wydanie 1.
- [7] E. L. Lehmann, *Fisher, Neyman, and the Creation of Classical Statistics*, Springer 2011, wydanie 1.
- [8] D. Salsburg, *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*, Henry Holt and Company 2002, wydanie 1.
- [9] P. Biecek, *Przewodnik po pakiecie R*, GiS 2018, wydanie 4.

OPIEKUN PRZEDMIOTU (IMIE, NAZWISKO, ADRES E-MAIL)

Andrzej Giniewicz (Andrzej.Giniewicz@pwr.edu.pl)