



Karta przedmiotu

| | | | | | | | |
|--|---|---|--|--------------|--|---|-------|
| Nazwa i kod przedmiotu | ANALIZA DANYCH Z R I PYTHON, PG_00060724 | | | | | | |
| Kierunek studiów | Analityka gospodarcza | | | | | | |
| Data rozpoczęcia studiów | październik 2023 r. | Rok akademicki realizacji przedmiotu | | | 2024/2025 | | |
| Poziom kształcenia | II stopnia | Grupa zajęć | | | Grupa zajęć fakultatywnych Grupa zajęć powiązanych z prowadzonymi badaniami naukowymi w dziedzinie nauki związanej z kierunkiem - profil ogólnoakademicki | | |
| Forma studiów | stacjonarne | Sposób realizacji | | | na uczelni | | |
| Rok studiów | 2 | Język wykładowy | | | polski | | |
| Semestr studiów | 3 | Liczba punktów ECTS | | | 4.0 | | |
| Profil kształcenia | ogólnoakademicki | Forma zaliczenia | | | zaliczenie | | |
| Jednostka prowadząca | Wydział Zarządzania i Ekonomii -> Katedra Statystyki i Ekonometrii | | | | | | |
| Imię i nazwisko wykładowcy (wykładowców) | Od odpowiedzialny za przedmiot | dr hab. Michał Pietrzak | | | | | |
| | Prowadzący zajęcia z przedmiotu | dr inż. Paulina Listewnik | | | | | |
| Formy zajęć i metody nauczania | Forma zajęć | Wykład | Ćwiczenia | Laboratorium | Projekt | Seminarium | RAZEM |
| | Liczba godzin zajęć | 0.0 | 0.0 | 45.0 | 0.0 | 0.0 | 45 |
| | W tym liczba godzin zajęć na odległość: 0.0 | | | | | | |
| Aktywność studenta i liczba godzin pracy | Aktywność studenta | Udział w zajęciach dydaktycznych, objętych planem studiów | Udział w konsultacjach | | Praca własna studenta | | RAZEM |
| | Liczba godzin pracy studenta | 45 | 8.0 | | 47.0 | | 100 |
| Cel przedmiotu | Proponuje nowatorskie rozwiązania złożonych i nieustrukturyzowanych problemów z wykorzystaniem nowoczesnych metod analitycznych, wspomagając swoją działalność wykorzystaniem technologii informatycznych | | | | | | |
| Efekty uczenia się przedmiotu | Efekt kierunkowy | | Efekt z przedmiotu | | | Sposób weryfikacji i oceny efektu | |
| | [K7_W03] demonstruje w pogłębionym stopniu przygotowanie w zakresie zastosowań metod analitycznych oraz technik formułowania i rozwiązywania problemów | | wykorzystuje możliwości wspomagania przez współczesne technologie informatyczne, w pogłębionej analizie danych ekonomiczno-społecznych | | | [SW1] Ocena wiedzy faktograficznej | |
| | [K7_U01] tworzy innowacyjne rozwiązania złożonych i nieustrukturyzowanych problemów uwzględniając zmienność otoczenia przez syntezę informacji pochodzących z wielu źródeł | | formułuje innowacyjne rozwiązania problemów ekonomiczno-społecznych, syntetyzując informacje z wielu źródeł- | | | [SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu | |

| | | | |
|---|---|---|-------------------------|
| Treści przedmiotu | <p>Wprowadzenie do języków R i Python. Podstawowe operacje Źródła danych. Import danych z różnych formatów w R vs Python Zmienne i typy danych w R vs Python (wektor, ramka danych, macierz, lista, tablice, zbiory, słowniki) Podstawowe funkcje - statystyki opisowe i matematyczne w R vs Python Podstawowe przetwarzanie danych (nowe zmienne, filtry, łączenie ramek: przekształcanie, dzielenie, łączenie) w R vs Python Brudne dane - brakujące obserwacje; duplikaty; wartości odstające; błędy formatowania. Pakiet naniar. Biblioteki Pythona Przetwarzanie danych przy użyciu Dplyr i Tidy. Przekształcanie danych w Pythonie Czyszczenie danych - wartości odstające i brakujące. Imputacje. Transformacje i dyskretyzacja zmiennych w Pythonie. Optymalny binning Grafika w R - podstawowa i zaawansowana graficzna prezentacja danych (pakiety: ggplot2; Lattice; Grid) vs Grafika w Pythonie (Matplotlib; Plotly itp.) Raportowanie analiz z R/Pythonem - wprowadzenie do R-Markdown i Quarto (notatnik; prezentacje - R i Powerpoint; HTML slidy; PDF beamer itp.). Shiny Apps dla Pythona Uczenie maszynowe - modele liniowe, uogólnione modele liniowe. Iteracyjne dopasowywanie modeli. Raportowanie Metoda k-najbliższych sąsiadów (KNN) Klasyfikacja i grupowanie. Liniowa analiza dyskryminacyjna; drzewa graficzne; regresja logistyczna Bagging i lasy losowe Metoda boosting. XGBoost</p> | | |
| Wymagania wstępne i dodatkowe | | | |
| Sposoby i kryteria oceniania osiągniętych efektów uczenia się | Sposób oceniania (składowe) | Próg zaliczeniowy | Składowa oceny końcowej |
| | Kolokwium | 60.0% | 50.0% |
| | Projekt - sprawozdania | 60.0% | 50.0% |
| Zalecana lista lektur | Podstawowa lista lektur | <p>Bruce Peter, Bruce Andrew, Gedeck Peter, Statystyka praktyczna w data science. 50 kluczowych zagadnień w językach R i Python, Helion, 2021 Chantal D. Larose, Daniel T. Larose, Data Science Using Python and R, Wiley, 2019 Rick J. Scavetta, Boyan Angelov, Python and R for the Modern Data Scientist, O'Reilly Media, 2021</p> | |
| | Uzupełniająca lista lektur | <p>Wes McKinney, Python w analizie danych. Przetwarzanie danych za pomocą pakietów Pandas i NumPy oraz środowiska IPython. Wydanie II. Helion, 2018 Marek Gagolewski, Maciej Bartoszek, Anna Cena, Przetwarzanie i analiza danych w języku Python, PWN, 2017 Ajay Ohri, Python for R Users: A Data Science Approach, Wiley, 2017 Hadley Wickham, Garrett Grolemund. R for Data Science, https://r4ds.had.co.nz J. Hathaway, Katie Larson, Python for Data Science, https://byuidatascience.github.io/python4ds/</p> | |
| | Adresy eZasobów | <p>Adresy na platformie eNauczanie: Analiza danych z R i Pythonem 2024 - Moodle ID: 39744 https://enauczenie.pg.edu.pl/moodle/course/view.php?id=39744</p> | |
| Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania | <p>Preprocessing danych w językach R oraz Python Raport eksploracyjnej analizy opisowej danych (DEA) w językach R oraz Python Wnioskowanie statystyczne z próby w językach R oraz Python</p> | | |
| Praktyki zawodowe w ramach przedmiotu | Nie dotyczy | | |

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.