

Karta przedmiotu

Nazwa i kod przedmiotu	Data Analysis and Visualisation in R, PG_00210490						
Kierunek studiów	Gospodarka przestrzenna (O)						
Data rozpoczęcia studiów	październik 2025 r.		Rok akademicki realizacji przedmiotu		2026/2027		
Poziom kształcenia	II stopnia		Grupa zajęć		Grupa zajęć obowiązkowych z zakresu kierunku studiów Grupa zajęć fakultatywnych		
Forma studiów	stacjonarne		Sposób realizacji		na uczelni		
Rok studiów	2		Język wykładowy		angielski		
Semestr studiów	3		Liczba punktów ECTS		2.0		
Profil kształcenia	ogólnoakademicki		Forma zaliczenia		zaliczenie		
Jednostka prowadząca							
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		dr Michał Marosz				
	Prowadzący zajęcia z przedmiotu						
Formy zajęć	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	30.0	0.0	0.0	0.0	0.0	30
	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	30		0.0		20.0	50
Cel przedmiotu	Celem przedmiotu jest rozwinięcie praktycznych umiejętności w zakresie importu, przetwarzania, analizy, interpretacji i wizualizacji danych z wykorzystaniem języka R oraz wybranych pakietów ekosystemu tidyverse . Studenci poznają podstawowe zasady pracy z danymi, przygotowywania ich do analiz, tworzenia czytelnych wizualizacji oraz prezentowania wyników w sposób przejrzysty, powtarzalny i zgodny z zasadami dobrej praktyki analitycznej.						
Efekty uczenia się przedmiotu	Efekt kierunkowy		Efekt z przedmiotu		Sposób weryfikacji i oceny efektu		
			Potrafi tworzyć i interpretować wizualizacje oraz proste raporty analityczne w R, wykorzystując pakiet ggplot2, kod, wyniki, wykresy i komentarz interpretacyjny do formułowania poprawnych wniosków.		[SK2] prezentacja/projekt/referat/ raport [SK4] test/egzamin - ustny lub pisemny		
			Potrafi importować, czyścić, przekształcać i analizować dane w R, w tym identyfikować braki danych, kontrolować typy zmiennych, wykonywać agregacje, łączenia tabel oraz podstawową eksploracyjną analizę danych.		[SK2] prezentacja/projekt/referat/ raport [SK4] test/egzamin - ustny lub pisemny		
			Zna podstawowe elementy środowiska pracy w R oraz rozumie zasady organizacji prostego projektu analitycznego, obejmującego skrypty, obiekty, ramki danych i powtarzalny przepływ pracy.		[SK2] prezentacja/projekt/referat/ raport [SK4] test/egzamin - ustny lub pisemny		

Treści przedmiotu	## Treści przedmiotu **Wprowadzenie do środowiska R i RStudio** * podstawowe elementy środowiska pracy, * skrypty, konsola, projekty R, * instalacja i ładowanie pakietów, * podstawowe typy obiektów w R. **Podstawy pracy z danymi w R** * ramki danych i tibbles, * struktura zbiorów danych, * typy zmiennych, * podstawowe operacje na obiektach i zmiennych. **Import danych z różnych źródeł** * import plików CSV i TXT, * import arkuszy kalkulacyjnych, * wykorzystanie przykładowych zbiorów danych dostępnych w pakietach R, * kontrola poprawności zaimportowanych danych. **Czyszczenie i przygotowanie danych do analizy** * identyfikacja braków danych, * kontrola i zmiana typów zmiennych, * wykrywanie podstawowych błędów w danych, * porządkowanie nazw kolumn i wartości zmiennych. **Przekształcanie danych z wykorzystaniem pakietów tidyverse** * filtrowanie obserwacji, * wybór i zmiana nazw zmiennych, * tworzenie nowych zmiennych, * grupowanie i agregacja danych, * sortowanie danych, * łączenie tabel. **Zmiana struktury danych** * dane w układzie szerokim i długim, * przekształcanie danych za pomocą funkcji <code>`pivot_longer()`</code> i <code>`pivot_wider()`</code>, * przygotowanie danych do analiz i wizualizacji. **Podstawowa eksploracyjna analiza danych** * opis rozkładów zmiennych, * podstawowe miary statystyki opisowej, * analiza zależności między zmiennymi, * porównywanie grup, * identyfikacja wartości odstających. **Wprowadzenie do wizualizacji danych w R** * rola wizualizacji w analizie danych, * dobór typu wykresu do charakteru danych, * podstawowe zasady czytelnej i poprawnej wizualizacji. **Tworzenie wykresów z wykorzystaniem pakietu ggplot2** * struktura wykresu w <code>`ggplot2`</code>, * wykresy punktowe, liniowe, kolumnowe i histogramy, * wykresy pudełkowe i gęstości, * mapowanie estetyk, * grupowanie danych na wykresach. **Dostosowywanie wyglądu wykresów** * skale osi, * kolory i palety barw, * etykiety, tytuły i legendy, * motywy graficzne, * poprawa czytelności wykresów. **Prezentacja zróżnicowania i niepewności danych** * przedziały ufności, * słupki błędów, * prezentacja rozrzutu i zmienności danych, * interpretacja niepewności na wykresach. **Interpretacja wyników analiz i wizualizacji** * formułowanie wniosków na podstawie wyników, * identyfikacja najważniejszych zależności i prawidłowości,
-------------------	---

	* unikanie błędów interpretacyjnych. 13. **Powtarzalność analiz** * organizacja projektu analitycznego, * dobre praktyki zapisu kodu, * dokumentowanie kolejnych etapów analizy, * podstawy pracy z raportem analitycznym. 14. **Przygotowanie prostego raportu analitycznego** * łączenie kodu, wyników, wykresów i komentarza, * struktura raportu analitycznego, * prezentacja wyników w sposób czytelny i powtarzalny. 15. **Samodzielna praca z dokumentacją i rozwiązywanie problemów** * korzystanie z pomocy w R, * interpretacja komunikatów błędów, * wyszukiwanie rozwiązań w dokumentacji i zasobach społeczności R, * dobre praktyki dalszego rozwijania umiejętności analitycznych.		
Wymagania wstępne i dodatkowe	Umiejętność posługiwania się komputerem PC na poziomie podstawowym. Znajomość podstaw z zakresu statystyki opisowej.		
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Kolokwium - praktyczne zastosowanie R w analizie i wizualizacji danych	51.0%	70.0%
	Projekt - analiza danych rzeczywistych (praca w grupie + prezentacja)	51.0%	30.0%
Zalecana lista lektur	Podstawowa lista lektur	1. Wickham, H., Çetinkaya-Rundel, M., Golemund, G. **R for Data Science. Import, Tidy, Transform, Visualize, and Model Data**. 2nd edition. O'Reilly Media. Pozycja podstawowa do nauki analizy danych w R z wykorzystaniem pakietów `tidyverse`. 2. Wickham, H., Navarro, D., Pedersen, T. L. **ggplot2: Elegant Graphics for Data Analysis**. Springer. Podstawowa pozycja dotycząca tworzenia wykresów w R z wykorzystaniem pakietu `ggplot2`. 3. Chang, W. **R Graphics Cookbook. Practical Recipes for Visualizing Data**. 2nd edition. O'Reilly Media. Praktyczne źródło przykładów i rozwiązań dotyczących wizualizacji danych w R. 4. Healy, K. **Data Visualization: A Practical Introduction**. Princeton University Press. Wprowadzenie do zasad skutecznej, czytelnej i odpowiedzialnej wizualizacji danych z wykorzystaniem R i `ggplot2`. 5. R Core Team. **An Introduction to R**. The R Project for Statistical Computing. Oficjalny podręcznik wprowadzający do języka R, jego składni, podstawowych struktur danych, grafiki i analiz statystycznych. 6. Golemund, G. **Hands-On Programming with R**. O'Reilly Media. Przystępne wprowadzenie do programowania w R, szczególnie przydatne dla osób rozpoczynających pracę z językiem.	
	Uzupełniająca lista lektur	1. Ismay, C., Kim, A. Y., Valdivia, M. **Statistical Inference via Data Science: A Modern Dive into R and the Tidyverse**. 2nd edition. CRC Press. Wprowadzenie do analizy danych, wizualizacji, regresji i podstaw wnioskowania statystycznego z wykorzystaniem R i `tidyverse`.	

	Adresy eZasobów	Podstawowe http://www.biecek.pl/R/ - Materiały udostępniane przez Przemysława Biecka https://education.rstudio.com/learn/ - RStudio-Education. Finding Your Way To R https://ggplot2-book.org/ - ggplot2: Elegant Graphics for Data Analysis https://r-graphics.org/ - R Graphics Cookbook, https://r4ds.hadley.nz/ - R4DataScience Oficjalna bezpłatna wersja online; bardzo dobra jako główny podręcznik kursu
--	-----------------	---

Przykładowe zadania praktyczne na kolokwium

Zadanie 1. Import i wstępna kontrola danych

Wczytaj plik `students\_results.csv`, który zawiera dane o wynikach studentów z kilku grup ćwiczeniowych.

Plik zawiera zmienne:

- * `student\_id` – identyfikator studenta,
- * `group` – grupa ćwiczeniowa,
- * `test\_1` – wynik pierwszego testu,
- * `test\_2` – wynik drugiego testu,
- * `project` – wynik projektu,
- * `attendance` – liczba obecności.

Wykonaj następujące operacje:

1. Wczytaj dane do R.
2. Sprawdź strukturę danych.
3. Sprawdź liczbę obserwacji i zmiennych.
4. Sprawdź, czy w danych występują braki danych.
5. Oblicz podstawowe statystyki opisowe dla zmiennych liczbowych.

Zadanie 2. Czyszczenie i przygotowanie danych

W zbiorze `students\_results.csv` niektóre nazwy grup zapisano niespójnie, np. `A`, `group A`, `Grupa A`, `B`, `group B`.

Wykonaj następujące operacje:

1. Ujednolić nazwy grup tak, aby miały postać `A`, `B`, `C`.
2. Zamień ewentualne braki danych w zmiennej `attendance` na wartość `0`.
3. Utwórz nową zmienną `final\_score`, która będzie średnią ważoną wyników:

- * `test\_1` – 30%,
- * `test\_2` – 30%,
- * `project` – 40%.

4. Utwórz nową zmienną `passed`, która przyjmuje wartość:

- * `"yes"` jeśli `final\_score >= 50`,
- * `"no"` jeśli `final\_score < 50`.

5. Wyświetl pierwsze 10 wierszy uporządkowanego zbioru danych.

Zadanie 3. Grupowanie, agregacja i interpretacja wyników

Na podstawie przygotowanego zbioru danych oblicz podsumowanie wyników dla każdej grupy ćwiczeniowej.

Dla każdej grupy oblicz:

1. liczbę studentów,
2. średni wynik końcowy,
3. medianę wyniku końcowego,
4. minimalny i maksymalny wynik końcowy,
5. odsetek studentów, którzy zaliczyli przedmiot.

Następnie:

6. uporządkuj grupy malejąco według średniego wyniku końcowego,
7. wskaż grupę z najwyższym średnim wynikiem,
8. sformułuj krótki komentarz interpretacyjny, składający się z 2–3 zdań.

Zadanie 4. Wizualizacja danych z wykorzystaniem ggplot2

Na podstawie danych `students\_results.csv` przygotuj trzy wykresy:

1. histogram zmiennej `final\_score`,
2. wykres pudełkowy pokazujący zróżnicowanie `final\_score` w grupach,
3. wykres kolumnowy przedstawiający odsetek studentów, którzy zaliczyli przedmiot w każdej grupie.

Każdy wykres powinien zawierać:

- * tytuł,
- * podpisane osie,
- * czytelną legendę, jeżeli jest potrzebna,
- * estetyczny motyw graficzny, np. `theme\_minimal()`.

	Na końcu wybierz jeden z wykresów i napisz krótką interpretację przedstawionych wyników. ## Przykładowy raport z analiz #### Zadanie 5. Prosty raport analityczny Przygotuj krótki raport analityczny w pliku <code>`.qmd`</code> lub <code>`.Rmd`</code>, wykorzystując dane <code>students_results.csv`</code>.   Raport powinien zawierać:       1. tytuł raportu,     2. krótki opis analizowanego zbioru danych,     3. kod importujący dane,     4. podstawowe podsumowanie danych,     5. tabelę z wynikami zagregowanymi według grup,     6. co najmniej dwa wykresy wykonane w <code>ggplot2`</code>, 7. krótki komentarz interpretacyjny do wyników, 8. końcowe wnioski w 3–5 zdaniach. Raport powinien być możliwy do wygenerowania jako dokument HTML lub PDF.
Praktyki zawodowe w ramach przedmiotu	Nie dotyczy

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.