



Uniwersytet Ekonomiczny w Poznaniu

Program szkolenia realizowanego w ramach projektu „Akademia kompetencji przyszłości”

Tytuł szkolenia	Analiza i wizualizacja danych w środowisku R	
Nazwa kompetencji/kwalifikacji uzyskiwanej po ukończeniu szkolenia	Analityk danych w języku R	
Wymiar godzinowy szkolenia	łącznie, w tym:	40 h
	a) w trybie face-to-face	
	b) w trybie online	40 h
Język szkolenia	polski	
Krótką charakterystyką szkolenia – wskazanie obszaru tematycznego i jego znaczenia	<p>Głównym celem szkolenia jest przedstawienie praktycznych możliwości wykorzystania języka R w obszarze przetwarzania danych. Uwaga skoncentrowana będzie na przybliżeniu informacji w zakresie najważniejszych typów danych (typy skalarne, napisy, listy, czynniki, wektory, macierze) oraz ramek danych (tworzenie ramek danych, importowanie ramek danych z plików i innych źródeł, filtrowanie, praca na kolumnach – zmiennych, przekształcanie zmiennych, sortowanie, obserwacje brakujące itp.). Ponadto omówione zostaną szczegółowo różne metody analizy danych statystycznych i sposoby ich wizualizacji w języku R. W trakcie szkolenia przybliżone zostaną w kontekście poruszanych treści zagadnienia związane z analizą i wizualizacją danych przede wszystkim w bibliotekach dplyr, sqldf, oraz graphics i ggplot2.</p>	
Adresaci szkolenia – charakterystyka grup docelowych	<p>Szkolenie jest przeznaczone dla szerokiego grona specjalistów z różnych branż, którzy chcą poszerzyć swoje umiejętności analityczne. Mogą to być analitycy biznesowi, specjaliści ds. marketingu, ekonomiści, informatycy, inżynierowie, urzędnicy, naukowcy, studenci kierunków ścisłych oraz osoby zmieniające ścieżkę kariery. Szczególnie cenni będą uczestnicy posiadający podstawową wiedzę informatyczną, logiczne myślenie i zainteresowanie analizą danych. Szkolenie jest dedykowane zarówno początkującym, którzy chcą poznać podstawy przetwarzania danych, jak i średniozaawansowanym specjalistom pragnącym usystematyzować i pogłębić swoją wiedzę z zakresu R oraz data science. Kluczowym atutem uczestników powinna być otwartość na naukę, chęć rozwiązywania problemów oraz gotowość do praktycznego zastosowania zdobytej wiedzy w codziennej pracy zawodowej.</p>	
Cele kształcenia	C1- poznanie środowiska oraz wybranych metod przetwarzania danych w języku R	
	C2 – poznanie podstawowych typów danych w języku R	
	C3- poznanie instrukcji sterujących i pisania własnych funkcji w języku R	
	C4- poznanie wybranych metod statystycznej analizy danych i ich wizualizacji w języku R	

**Efekty uczenia się dla szkolenia oraz kryteria i metody ich weryfikacji**

Kod efektu dla szkolenia	Efekty uczenia się	Kryteria weryfikacji ¹ osiągnięcia poszczególnych efektów uczenia się	Metody weryfikacji osiągnięcia poszczególnych efektów uczenia się
Wiedza			
W1	Uczestnik zna i rozumie możliwości wykorzystania języka R w obszarze przetwarzania danych	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
W2	Uczestnik zna i rozumie podstawowe typy danych, instrukcje sterujące i funkcje w języku R	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
W3	Uczestnik zna i rozumie możliwości praktycznego wykorzystania języka R w obszarze statystycznej analizy danych	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
W4	Uczestnik zna i rozumie możliwości praktycznego wykorzystania języka R w obszarze wizualizacji danych	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
Umiejętności			
U1	Uczestnik potrafi przetwarzać dane na różnych ich typach	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
U2	Uczestnik potrafi samodzielnie stosować instrukcje sterujące i pisać własne funkcje w języku R	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia	obserwacja pracy podczas zajęć, analiza pisanych

¹ „Kryteria weryfikacji to działania, które powinna wykonać osoba w trakcie weryfikacji, aby udowodnić, że ma wymagane efekty uczenia się”.



Kod efektu dla szkolenia	Efekty uczenia się	Kryteria weryfikacji ¹ osiągnięcia poszczególnych efektów uczenia się	Metody weryfikacji osiągnięcia poszczególnych efektów uczenia się
		uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	kodów, końcowy sprawdzian testowy
U3	Uczestnik potrafi wykorzystać język R w obszarze statystycznej analizy danych i ich wizualizacji	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
U4	Uczestnik potrafi interpretować wyniki przeprowadzonych analiz	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
Kompetencje społeczne			
K1	Uczestnik jest świadom znaczenia przydatności języków programowania w analizie zjawisk społeczno-gospodarczych	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
K2	Uczestnik jest świadom roli metod przetwarzania danych, ich analizy i wizualizacji w opisie zjawisk społeczno-gospodarczych	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy
K3	Uczestnik jest gotów na samodzielne uzupełnianie posiadanej wiedzy w zakresie metod przetwarzania danych, ich analizy i wizualizacji	tworzy samodzielnie kod, stosuje proste i złożone metody przetwarzania danych, omawia uzyskane wyniki analiz, samodzielnie poszerza wiedze na temat programowania oraz przetwarzania danych	obserwacja pracy podczas zajęć, analiza pisanych kodów, końcowy sprawdzian testowy



Spis omawianych zagadnień

Kod	Tytuł zagadnienia	Liczba godzin	Realizowane cele kształcenia /kody celów kształcenia/
Z1	Środowisko RStudio w usłudze chmurowej	1	C1
Z2	Język Markdown i RMarkdown	1	C1
Z3	Typy skalarne	2	C1, C2, C4
Z4	Czynniki i listy	2	C1, C2, C4
Z5	Wektory	4	C1, C2, C4
Z6	Macierze	2	C1, C2, C4
Z7	Ramki danych	8	C1, C2, C4
Z8	Przetwarzanie danych w pakiecie dplyr	4	C1, C2, C4
Z9	Wprowadzenie do SQL w R	4	C1, C2, C4
Z10	Instrukcje sterujące	4	C1, C3, C4
Z11	Funkcje w języku R	4	C1, C3, C4
Z12	Wizualizacja danych w pakiecie ggplot2	4	C1, C4

Szczegółowy opis omawianych zagadnień

Kod zag.	Szczegółowy opis zagadnienia	Nabywane efekty uczenia się /kod efektu uczenia się/
Z1	Wprowadzenie do RStudio, omówienie najważniejszych okien	W1, U1, U4, K1
Z2	Omówienie przeznaczenia języka znaczników Markdown, formatowanie tekstu w języku Markdown, nagłówki, listy numerowane, listy punktowane, linki i obrazy, cytaty oraz kod, tabele i wzory w języku Markdown i RMarkdown	W1, U1, U4, K1



Kod zag.	Szczegółowy opis zagadnienia	Nabywane efekty uczenia się /kod efektu uczenia się/
Z3	Omówienia podstawowych typów skalarnych, podstawowe operatory arytmetyczne, podstawowe operatory relacyjne, operacje na liczbach całkowitych, operacje na liczbach zmiennoprzecinkowych, operacje na liczbach zespolonych, konwersje między typami	W1, W2, W3, U1, U3, U4, K1, K2, K3
Z4	Czynniki i ich zastosowania, listy oraz ich zastosowania, tworzenie list, dostęp do elementów listy, modyfikacja elementów listy, dodawanie i usuwanie elementów listy, funkcje agregujące na listach	W1, W2, W3, U1, U3, U4, K1, K2, K3
Z5	Znaczenie i zastosowania wektorów, tworzenie wektorów, indeksowanie i wycinanie, łączenie wektorów, operacje matematyczne na wektorach, agregacja i podstawowe statystyki	W1, W2, W3, U1, U3, U4, K1, K2, K3
Z6	Znaczenie i zastosowania macierzy, tworzenie macierzy, indeksowanie i wycinanie, łączenie macierzy, operacje matematyczne na macierzach, agregacja i podstawowe statystyki	W1, W2, W3, U1, U3, U4, K1, K2, K3
Z7	Tworzenie ramek danych na podstawie wektorów i macierzy, wczytywanie danych z różnych formatów do ramek danych, wybór kolumn i wierszy, dodawanie kolumn, usuwanie kolumn, dodawanie wierszy, zmiana typu danych w ramach danych, agregacja i podsumowania, operacje grupowania, sortowanie danych, zapis danych do różnych plików, operacje na danych	W1, W2, W3, U1, U3, U4, K1, K2, K3
Z8	Przetwarzanie danych z wykorzystaniem pakietu dplyr, wybór wierszy i kolumn, sortowanie danych, tworzenie nowych zmiennych, agregacja i wybrane statystyki opisowe	W1, W2, W3, U1, U3, U4, K1, K2, K3
Z9	Omówienie języka SQL, omówienie funkcji sqldf() w R, główne parametry funkcji sqldf(), tworzenie różnych zapytań, agregacja danych	W1, W2, W3, U1, U3, U4, K1, K2, K3
Z10	Rodzaje instrukcji sterujących, instrukcje warunkowe, składnia instrukcji warunkowej, pętla while, pętla for, modyfikowanie elementów pętli, obsługa wyjątków	W1, W2, W3, U1, U2, U3, U4, K1, K2, K3
Z11	Definicja funkcji, parametry i argumenty, tworzenie własnych funkcji	W1, W2, W3, U1, U2, U3, U4, K1, K2, K3
Z12	Podstawowe biblioteki do wizualizacji danych (graphics, ggplot2, plotly), rysowanie punktów i łamanych, wielokąty i wypełnienia, adnotacje tekstowe, różne typy wykresów (punktowe, kołowe, histogramy, pudełkowe, słupkowe)	W1, W2, W3, W4, U1, U3, U4, K1, K2, K3

Metody nauczania stosowane przy poszczególnych zagadnieniach

Kod zag.	Omówienie metod nauczania
Z1	Wykład z prezentacją multimedialną, rozwiązywanie zadań
Z2	Wykład z prezentacją multimedialną, rozwiązywanie zadań
Z3	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z4	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z5	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków



Kod zag.	Omówienie metod nauczania
Z6	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z7	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z8	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z9	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z10	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z11	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków
Z12	Wykład z prezentacją multimedialną, rozwiązywanie zadań, analiza przypadków

Wymagania wstępne:

- znajomość podstawowych miar statystycznych: średnia, odchylenie standardowe, kwartyle itp.
- wcześniejsza znajomość języka R nie jest wymagana

Warunki ukończenia szkolenia (czyli zdobycia kompetencji/kwalifikacji):

- zaliczenie końcowego sprawdzianu testowego z wykorzystaniem języka R (test i jego zasady omówione będą podczas ostatnich zajęć – sam test z rozwiązaniami będzie musiał zostać wysłany wraz z kodem języka R prowadzącemu zajęcia na wskazany adres mailowy w ciągu tygodnia od momentu jego przekazania słuchaczom). Każdemu zadaniu w teście przypisana będzie odpowiednia liczba punktów możliwych do uzyskania przez słuchacza za jego poprawne rozwiązanie. Próg zaliczenia: 60% łącznej możliwej liczby punktów do zdobycia.

Materiały dydaktyczne, jakie zostaną przekazane uczestnikom szkolenia:

- prezentacja na temat języka R
- zestawy zadań do rozwiązywania na zajęciach i do samodzielnego wykonania
- zbiory danych



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Certyfikacja:

***** *informacje nt. dodatkowych certyfikatów (oprócz standardowego certyfikatu ukończenia szkolenia), o ile ich uzyskanie jest przewidziane programem szkolenia*

- standardowy certyfikat ukończenia szkolenia

Autor programu szkolenia:

dr hab. Marcin Szymkowiak, prof. UEP