



UNIVERSITY
OF WARSAW

Faculty of Modern Languages
School for International Science

R M

Research Methodology



R 101: porządkowanie danych, analiza i wizualizacja*

***Kurs możliwy do przeprowadzenia w języku angielskim.**

Wydział Neofilologii UW zaprasza pracowników i doktorantów naszego Uniwersytetu na kurs wprowadzający do użycia programu R: **R 101: porządkowanie danych, analiza i wizualizacja**. R to język programowania oraz program coraz częściej używany do analizy danych, między innymi dlatego, że jest darmowy, łatwo dostępny, a w Internecie można znaleźć wiele podpowiedzi i wskazówek obsługi. Na tym kursie nauczymy się używać R do wykonywania podstawowych czynności podczas analizy danych: porządkowania, analizy i wizualizacji danych.

Kurs adresowany jest do osób, które:

- chcą rozwijać i doskonalić warsztat analizy i wizualizacji danych,
- chciałyby rozpocząć swoją przygodę z R,
- znają podstawowe czynności związane z analizą danych, np. filtrowanie danych,
- przeprowadzały już (proste) analizy statystyczne, takie jak porównanie grup, korelacje.

Kurs *R 101: porządkowanie danych, analiza i wizualizacja* obejmuje 30 godzin zajęć w formie praktycznych ćwiczeń: uczestnicy będą pracować w programie R od pierwszych zajęć, na danych udostępnionych w środowisku R lub danych prowadzącej. Kurs podzielony jest na 3 moduły poświęcone: (1) porządkowaniu danych (import danych, czyszczenie, podzbiory, mutacja, czyli wykorzystanie istniejących zmiennych do tworzenia nowych); (2) analizie danych (statystyki opisowe, porównania grupowe, korelacje); (3) wizualizacji danych przy użyciu ggplot2 (wykresy pudełkowe, wykresy rozrzutu, wykresy liniowe, dostosowywanie i eksportowanie wykresów). Uczestnicy otrzymają także 4 zadania domowe (nieoceniane), które będą wymagały od nich wykorzystania dotychczas zdobytej wiedzy. Na koniec zajęć każdy uczestnik dostanie projekt (dane oraz serię zadań) do wykonania w R.

Zastrzeżenie: Kurs *R 101: porządkowanie danych, analiza i wizualizacja* nie jest kursem programowania w języku R (tj. nie uczy instrukcji warunkowych, pętli, pisania własnych funkcji). Nie są to również typowe zajęcia ze statystyki (tj. kurs nie wskaże, który test statystyczny najlepiej pasuje do danych). Jego celem jest usprawnienie codziennej pracy badawczej.

CELE KURSU:

Uczestnicy kursu *R 101: porządkowanie danych, analiza i wizualizacja* dowiedzą się:

- jak używać programu i języka R do porządkowania, analizy i wizualizacji danych;
- jak szukać rozwiązań i odpowiedzi dotyczących R w internecie;
- jak zadbać o dalsze rozwijanie swoich umiejętności w R.

EFEKTY KSZTAŁCENIA:

Po ukończeniu kursu *R 101: porządkowanie danych, analiza i wizualizacja* **uczestnik będzie znał:**

- środowisko R,
- typowe sposoby pozyskiwania informacji i odpowiedzi na pytania dot. R.

Uczestnik będzie potrafił:

- importować dane, filtrować, tworzyć podzbiory, mutować dane (czyli wykorzystywać istniejące zmienne do tworzenia nowych),
- przeprowadzić podstawowe analizy danych: statystyki opisowe, porównania grup, korelacje,
- opracować wizualizację danych: stworzyć, dostosować i eksportować wykres, zastosować nabytą wiedzę i umiejętności we własnej praktyce naukowej.

Uczestnik będzie doceniał:

- rolę programów takich jak R w procesie analizy danych.

SYLABUS (30h)

Moduł I:

Porządkowanie danych

10
godzin

TREŚĆ:

- Przedstawienie R, R Studio; Dlaczego R?
- Gdzie szukać pomocy: Dokumentacja R, R-bloggers, Stackoverflow, itp.
- Import danych z Excela, SPSS, plików tekstowych
- Przygotowanie pliku do użycia w R: format tidy verse (kolumny to zmienne, wiersze to obserwacje)
- Filtrowanie, selekcja, mutowanie zmiennych
- Formaty szerokie i długie

Moduł II:

Analiza danych

10
godzin

TREŚĆ

- Podstawowa eksploracja danych: statystyki opisowe, histogramy
- Porównanie grup (2 lub więcej; between subjects i within subjects, ewentualnie: parametryczne i nieparametryczne)
- Korelacje, matryce korelacji
- Używanie RMarkdown i/lub RNotebooks do raportowania wyników

Moduł III:

Wizualizacja danych

10
godzin

TREŚĆ:

- Wykresy pudełkowe (boxplots)
- Wykresy rozrzutu (scatterplots)
- Wykresy liniowe (linegraphs)
- Dostosowywanie wykresów do własnych potrzeb
(np. rozmiary, kolory, etykiety osi)
- Tworzenie grup i paneli wykresów, eksport wykresów

KONTAKT

School for International Science
Wydział Neofilologii | Uniwersytet Warszawski
www.sis.uw.edu.pl | mail: sis@uw.edu.pl