

R як середовище обробки даних соціологічних досліджень

Автор курсу - кандидат соціологічних наук, науковий співробітник Інституту соціології НАН України Сергій Дембіцький.

Рівень - для початківців.

Чому цей курс є актуальним

R є мовою програмування та комплексним середовищем для статистичної обробки даних. Можливості R перевищують можливості SPSS. Крім того R є цілком безкоштовним продуктом, який постійно розвивається силами великого об'єднання науковців зі всього світу. Втім робота в R передбачає набуття додаткових навиків (які часто є незвичними для соціологів) управління даними - роботи з синтаксисом. Власне пропонований авторський курс і ставить на меті сформувати у слухачів навик управління даними шляхом роботи з синтаксисом в R. Курс носить практичну спрямованість - весь пропонований матеріал спирається на спільному виконанні практичних завдань.

Наприкінці окрема увага приділяється методам статистичного аналізу за допомогою бібліотеки «Rcmdr», яка призначена для створення графічного інтерфейсу користувача. Її функціонал є неменьшим ніж функціонал відомого в Україні програмного продукту статистичної обробки даних «OCA for Windows».

Треба пам'ятати - можливості R значно перевищують потреби пересічного соціолога, але тільки тому, що йому невідомі ці можливості. Звідси філософія роботи в R: «Нові можливості => нові потреби => нові результати!».

Модуль 1 (2 дні по 3 пари) Основи управління даними в R

Основна мета - сформувати у слухачів основні навички підготовки соціологічних масивів до аналізу засобами R: створенням масивів «з нуля», імпортом масивів, перекодуванням та обчисленням змінних, особливостями числових та категоріальних змінних.

Теми

1. Типи та властивості об'єктів в R. Створення масивів даних.
2. Управління даними: змінні-фактори, обчислення та перекодування.
3. Вбудовані функції в R.

Модуль 2 (2 дні по 3 пари) Основи програмування в R: перший рівень

Основна мета - сформувати у слухачів навички роботи з базовими конструкціями мов програмування, що дозволяють автоматизувати та оптимізувати обробку й аналіз соціологічних даних.

Теми

4. Основи програмування: цикли.
5. Основи програмування: умовні оператори.
6. Основи програмування: використання списків.

Модуль 3 (2 дні по 3 пари) Основи програмування в R: другий рівень

Основна мета - сформувати у слухачів навички побудови власних функцій, які забезпечують вирішення нестандартних аналітичних завдань, а також продемонструвати можливості графічного інтерфейсу користувача «R Commander», який полегшує статистичний аналіз даних засобами R.

Теми

7. Основи програмування: авторські функції.
8. Основи програмування: рекурсія та зміна глобальних об'єктів.
9. Вступ до аналізу даних.