	U-LAS70 10002 SE50						
< >	ILAS Seminar-E2 :Programming for data analysis ILAS Seminar-E2 :Programming for data analysis				VANDENBON Alexis		
			2		1		
	2025	**(1)**	15 (15)				
	5			23			
	Software-related / Programming / R / Data analysis / Statistics						

[   ]

R programming language is a useful environment for statistical data analysis and machine learning. The R language is widely used in many fields of science, for data processing, analysis, and visualization. In this course, I will introduce basic R programming techniques. Using example applications, I will illustrate how to use R to process and manipulate data, to write your own functions, to perform statistical tests, and to make figures.

[   ]

Students will learn the basic features of the R language for data manipulation, computation, and visualization. They will learn how to write your own code and functions, and how to use publicly available packages. Example applications introduced during the course will give students enough experience to use R for their own analysis.

[   ]

Lecture 1: Introduction to R. We will introduce R, its main features, and advantages and disadvantages. Using R interactively we will introduce some simple data types and commands.
Lectures 2-3. Simple manipulations, numbers and vectors. In this session, we will continue introducing simple operations. We will also discuss vectors, how to access their elements, and how to manipulate them.
Lecture 4: Inspecting variables and the workspace. We will discuss the properties of different classes of variables, and how to manipulate variables and the workspace.
Lectures 5-6: We will cover how to make vectors, arrays and matrices, and how to apply commands on them. We will introduce ways to manipulate arrays and matrices, and how to store and access data in them.
Lecture 7: Lists and data frames. We will introduce lists and data frames, and their basic commands and features.
Lecture 8: Environments and functions. So far we have only used pre-defined functions. In these two lectures we will discuss how to write your own functions for manipulating and processing various types of data.
Lecture 9: Flow control and loops. We will introduce ways how to execute commands only when some conditions are met (if statements), and how to execute operations repeatedly (various types of loops).
Lecture 10: Packages. Apart from pre-installed functions, there are thousands of libraries and packages publicly available. Here we will discuss how to find such packages in the " Comprehensive R Archive Network" (CRAN), how to install them, find documentation, and use them.
Lecture 11: Getting data and cleaning data. We will discuss several ways of reading data from files, cleaning data, and how to save data in files.
Lecture 12: Data visualization. We will introduce 3 big approaches for making various types of plots and figures in R.

Lecture 13: Statistical tests and probability distributions. R is particularly useful for statistical analysis of data. We will introduce commands related to probability distributions, and commands for applying various widely used statistical tests.
Lecture 14. Review of course material.
Lecture 15: Feedback

[   ]

[   ]

Grading: Attendance and active participation (20%) and small quizzes at the end of lectures (80%).

[   ]

Richard Cotton　　Learning R: A Step-by-Step Function Guide to Data Analysis (first edition)　　　O'Reilly Media　ISBN:978-1449357108　　The course lectures will roughly follow the content of this textbook. It will be supplemented with additional material, including an introduction to R available on the CRAN website (https://cran.r-project.org/manuals.html).

[   ]

The course is based on the content of the textbook "Learning R: A Step-by-Step Function Guide to Data Analysis", but it is not necessary to buy the book.

[   ]

It is strongly recommended to bring a laptop to the class.
No fixed office hours. Students are requested to make appointments directly or by email.

[   ]