1.3.4. Data Analysis in R – Beginners Course (Cologne University of Applied Sciences)

Name of Module	Data Analysis in R - Beginners Course
Short description	In this course you will learn a programming language and how to work with large amounts of data. Not only will these skills increase the possibilities for what you can achieve in your studies, they are increasingly becoming a requirement to gain employment in many fields. They provide you numerous opportunities for the modern day professional work. This course teaches how to use The R Project for Statistical Computing (commonly known as "R") for data analysis, focusing on the processing and analysis of spatial and temporal datasets. The intensive course starts at a beginner level and moves to an intermediate level. Please note that the course uses examples and data analysis techniques in the fields of **climate, geography** and **hydrology,** and it is therefore recommend that students in master's courses related to these topics attend.
Name of Programme	
Name of University	
Name of Lecturer	Dr. Oscar Manuel Baez Villanueva
Responsible University lecturer	

Credit Points	**SWS**	**Attendance (h)**	**Self-study (h)**	**Total workload (h)**
None	-	17.5	22.5	40

Start & end dates	**Timeslot:**
December 2023	7 interactive lectures and 2 sessions allocated for students to work on exercises.
Registration until	**Number of possible AGEP participants**
to be announced on the AGEP website	30

Content and goals of qualification	Objective: for students to obtain and implement the skills to undertake geospatial data analysis using the R Project for Statistical Computing. **Module 1: Introduction to R part I** 1. What is R? 2. R and Rstudio 3. R commands 4. Assignment, vectors and sequences 5. Missing values 6. Index vectors 7. Objects and their modes and attributes 8. Arrays and matrices 9. Lists and data frames 10. Scripts and packages 11. Basic functions 12. Arithmetic operators **Module 2: Introduction to R part II**

	1. Importing and exporting data 2. Basic plotting 3. Relational operators 4. Iterative processes **Module 3: Raster files and spatial data** 1. Importing and plotting vector data 2. Extract specific spatial information using vector data 3. Importing gridded datasets with different formats 4. Manipulation of gridded datasets 5. Exporting gridded datasets in different formats 6. Manipulation of values in gridded data **Module 4: Satial data processing** 1. Data organisation 2. Accessing and importing particular files 3. Pre-processing data 4. Aggregating data 5. Generation of time series from gridded datasets 6. Extra Considerations **Module 5: Spatial and temporal statistics** 1. Raster statistics a. Minimum, maximum, mean, standard deviation, sum b. Frequency of cells in a raster c. Writing derived values as a time series 2. Example: Raster statistics over an area a. Calculating mean P and ETa over an area b. Analysing P minus ETa patterns **Module 6: Automating downloading** 1. Web crawling with R 2. Iterative functions 3. Downloading Products trough ftps 4. Downloading Products trough https
Preconditions	Basic knowledge of statistics
Teaching Methods	Online, via Zoom
Lesson format (online/face-to-face)	The format will be online and includes: 1. Lectures 2. Exercises 3. Data and scripts to reproduce examples and solve the exercises 4. Questions and Answer (Q&A) sessions
Assessment	None
Language	English
registration	www.agep-info.de
Certificate	Provided by AGEP / TH Köln / DAAD