# Data Analysis in R, Cologne University of Applied Sciences

| Name of Module | Data Analysis in R |
|---|---|
| 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.<br>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 | Natural Resources Management |
| Name of University | Cologne University of Applied Sciences |
| Name of Lecturer | Oscar Manuel Baez Villanueva & Ian McNamara |
| Responsible University lecturer | Prof. Lars Ribbe |

| Credit Points | SWS | Attendance (h) | Self-study (h) | Total workload (h) |
|---|---|---|---|---|
| no | - | 25 | 25 | 50 |

| Start & end dates, WS | Timeslot: |
|---|---|
| **28.9.2020 – 9.10.2020** | 2 weeks, 2.5 h daily, Mondays to Fridays, 9:00 - 11:30 |

| Registration until | Number of possible AGEP participants |
|---|---|
| September 11th 2020 | 50 |

| 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.<br><br>**Module 1: Introduction to R part I**<br>1. What is R?<br>2. Scripts and packages<br>3. Data in R<br>4. Mode and length<br>5. Some basic functions<br>6. Arithmetic operators<br>7. Matrices<br>8. Lists<br>9. Data frames<br><br>**Module 2: Introduction to R part II**<br>1. Importing data<br>2. Writing data<br>3. Basic plotting<br>4. Relational operators<br>5. Loops<br>6. |

**Module 3: Raster files and spatial data**
1. Reading and plotting shapefiles
2. Extracting polygons from shapefiles
3. Reading rasters
4. Stacking, cropping, masking and resampling rasters
5. Writing rasters
6. Isolating cells with particular attributes

**Module 4: Data processing**
1. Data organisation
2. Accessing and loading particular files
3. Pre-processing data
   a. Example: CHIRPS precipitation raw data
4. Aggregating data
   a. Example: converting monthly to annual data
5. Extra Considerations
   a. Temporary files
   b. Computer cores

**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 (P minus ETa)
   a. Calculating mean P and ETa over an area
   b. Analysing P minus ETa patterns

**Module 6: Performance Indicators (Streamflow and Point-to Pixel Evaluation)**
1. Introduction to performance indicators
   a. Common performance indicators
   b. Working with NA values
2. Comparing time series in R
3. Example: Evaluating streamflow using performance indicators
4. Example: Point-to-Pixel evaluation

**Module 7: Automating downloading**
1. The apply, lapply, sapply and mapply functions
2. Reading netcdf Files
   a. Example: ERA5 Data
3. Downloading Products with ftps
   a. Example: CHIRPSv2 (Monthly)
4. Downloading MODIS products
5. Packages in R
6. Functions in R

| | |
|---|---|
| **Preconditions** | Basic knowledge of statistics |
| **Teaching Methods** | Online, via Zoom |
| **Lesson format (online/face-to-face)** | The format will be online and includes:<br>1. Lectures<br>2. Exercises<br>3. Data and scripts to reproduce examples and solve the exercises<br>4. Questions and Answer (Q&A) sessions |
| **Assessment** | None |
| **Language** | English |
| **registration** | www.agep-info.de |
| **Certificate** | Provided by AGEP / TH Köln / DAAD |