

# Introduction to R for microbial ecologists

## Course outline

### **Module 1: Introduction to R Syntax and Data Structures**

Module 1: Introduction to R Syntax and Data Structures is an introductory course designed to provide microbial ecologists with the necessary skills to use the R programming language. This module covers the basics of R syntax, data structures, and data manipulation. It also introduces the use of packages and functions to work with data. By the end of this module, participants will have a solid understanding of the fundamentals of R and be able to use it to analyze and visualize microbial ecology data.

#### ***Lessons***

- Overview of R Syntax and Data Structures
- Working with Vectors in R
- Working with Matrices in R
- Working with Lists in R
- Working with Data Frames in R
- Working with Factors in R
- Working with Strings in R
- Working with Dates and Times in R
- Working with Missing Values in R
- . Working with Factors in R for Microbial Ecologists

#### **After completing this module, students will be able to:**

- Understand the basic syntax of the R programming language.
- Create and manipulate data structures such as vectors, matrices, and data frames.
- Use functions to perform basic operations on data.
- Visualize data using basic plotting functions.

### **Module 2: Exploring and Manipulating Data in R**

Module 2 of the Introduction to R for Microbial Ecologists course provides an introduction to the basics of exploring and manipulating data in R. Students will learn how to import data into R, explore and summarize data, and create visualizations. They will also learn how to use basic data manipulation techniques such as subsetting, merging, and reshaping. Finally, they will learn how to use the dplyr package to manipulate data.

## ***Lessons***

- Introduction to Data Structures in R
- Working with Data Frames in R
- Subsetting and Sorting Data in R
- Exploring Data with Descriptive Statistics in R
- Visualizing Data with Plots in R
- Working with Factors in R
- Joining and Merging Data in R
- Reshaping Data with tidyr and reshape2 in R
- Working with Dates and Times in R
- . Writing Functions in R

### **After completing this module, students will be able to:**

- Understand the basics of data manipulation and analysis in R
- Be able to import, export, and manipulate data in R
- Be able to create basic visualizations of data in R
- Be able to perform basic statistical tests in R to analyze microbial ecology data

## **Module 3: Visualizing Data in R**

Module 3: Visualizing Data in R is an introduction to the basics of data visualization in R. It covers topics such as creating basic plots, customizing plots, and using packages to create more complex visualizations. This module is designed to help microbial ecologists gain the skills necessary to effectively visualize their data in R.

## ***Lessons***

- Introduction to ggplot2
- Creating Bar Plots and Histograms
- Working with Scatter Plots
- Creating Line Graphs
- Working with Box Plots
- Creating Heat Maps
- Working with Bubble Plots
- Creating Pie Charts
- Working with Network Graphs
- . Creating 3D Plots

### **After completing this module, students will be able to:**

- Create visualizations of microbial data using R packages such as ggplot2 and lattice.
- Understand the principles of data visualization and how to apply them to microbial data.
- Utilize the R programming language to manipulate and visualize microbial data.
- Interpret and communicate the results of data visualizations to other scientists.

## Module 4: Statistical Analysis in R

Module 4: Statistical Analysis in R is an introduction to the fundamentals of statistical analysis in the R programming language. This module covers topics such as descriptive statistics, linear regression, ANOVA, and other advanced topics. It is designed to provide a comprehensive overview of the statistical analysis capabilities of R for microbial ecologists.

### **Lessons**

- Exploring Data with Descriptive Statistics
- Visualizing Data with Graphs
- Testing for Normality
- Correlation and Regression Analysis
- ANOVA and Linear Models
- Non-parametric Tests
- Multivariate Analysis
- Time Series Analysis
- Survival Analysis
- . Machine Learning Algorithms in R

### **After completing this module, students will be able to:**

- Understand the fundamentals of statistical analysis in R, including basic concepts such as probability distributions, hypothesis testing, and linear regression.
- Utilize R to perform basic statistical tests such as t-tests, ANOVA, and chi-square tests.
- Interpret the results of statistical tests and create meaningful visualizations of the data.
- Apply statistical analysis to microbial ecology data to answer research questions.

## Module 5: Working with Microbial Data in R

Module 5 of the Introduction to R for Microbial Ecologists course provides an introduction to working with microbial data in R. It covers topics such as importing data, data manipulation, data visualization, and statistical analysis. This module is designed to give students the skills they need to analyze microbial data in R.

### **Lessons**

- Introduction to R for microbial ecologists
- Exploring microbial data in R
- Visualizing microbial data in R
- Analyzing microbial data in R
- Working with microbial data frames in R
- Working with microbial data sets in R
- Working with microbial data tables in R
- Working with microbial data matrices in R
- Working with microbial data objects in R
- . Working with microbial data structures in R

- . Working with microbial data analysis in R
- . Working with microbial data manipulation in R
- . Working with microbial data mining in R
- . Working with microbial data wrangling in R
- . Working with microbial data visualization in R
- . Working with microbial data exploration in R
- . Working with microbial data transformation in R
- . Working with microbial data cleaning in R
- . Working with microbial data integration in R
- . Working with microbial data aggregation in R

### **After completing this module, students will be able to:**

- Understand the basics of working with microbial data in R, including data manipulation, visualization, and analysis.
- Be able to use R to explore microbial data, including data from 16S rRNA sequencing, metagenomics, and other sequencing technologies.
- Be able to use R to perform basic statistical tests on microbial data, such as ANOVA, t-tests, and linear regression.
- Be able to use R to create publication-quality figures for microbial data.

## **Module 6: Modeling Microbial Communities in R**

Module 6: Modeling Microbial Communities in R is an introduction to the use of R for microbial ecologists. It covers topics such as data manipulation, data visualization, and statistical modeling of microbial communities. It also provides an overview of the different types of models available and how to use them to analyze microbial community data.

### **Lessons**

- Introduction to the tidyverse for microbial ecology
- Exploring microbial community data with ggplot2
- Visualizing microbial community data with heatmaps
- Analyzing microbial community data with ordination
- Modeling microbial community data with linear and non-linear regression
- Modeling microbial community data with machine learning algorithms
- Modeling microbial community data with random forests
- Modeling microbial community data with support vector machines
- Modeling microbial community data with artificial neural networks
- . Modeling microbial community data with Bayesian networks

### **After completing this module, students will be able to:**

- Understand the principles of microbial ecology and how to apply them to data analysis.
- Utilize R to analyze microbial community data.
- Interpret and visualize microbial community data using R.
- Develop models to predict microbial community dynamics.

## Module 7: Advanced Topics in R for Microbial Ecologists

Module 7: Advanced Topics in R for Microbial Ecologists is an in-depth exploration of the more complex aspects of R programming for microbial ecologists. This module covers topics such as data manipulation, plotting, and statistical analysis. It also covers topics such as machine learning, text mining, and web scraping. This module is designed to give microbial ecologists the skills they need to effectively analyze and interpret their data.

### *Lessons*

- Exploring Microbial Communities with R
- Visualizing Microbial Data with R
- Analyzing Microbial Diversity with R
- Modeling Microbial Dynamics with R
- Working with High-Throughput Sequencing Data in R
- Working with Metagenomic Data in R
- Working with Microbial Genome Data in R
- Working with Microbial Transcriptomic Data in R
- Working with Microbial Proteomic Data in R
- . Working with Microbial Metabolomic Data in R
- . Working with Microbial Phenotypic Data in R
- . Working with Microbial Population Data in R
- . Working with Microbial Interaction Data in R
- . Working with Microbial Pathway Data in R
- . Working with Microbial Network Data in R
- . Working with Microbial Ecological Data in R
- . Working with Microbial Evolutionary Data in R
- . Working with Microbial Genomics Data in R
- . Working with Microbial Metagenomics Data in R
- . Working with Microbial Metatranscriptomics Data in R

### **After completing this module, students will be able to:**

- Understand the principles of microbial ecology and how to apply them in R.
- Develop an understanding of the different types of data analysis techniques available in R and how to use them to analyze microbial data.
- Utilize advanced R packages and functions to analyze microbial data.
- Create visualizations of microbial data using R packages and functions.