Programming and Data Analysis with R

Dr. Anja Mühlemann

Variables

- Variables allow you to store values and reuse them multiple times in your code. This helps avoid rewriting the same values repeatedly.
- Using self-explanatory names like height, age, temperature improves code readability, making it easier to understand what the code does.
- By defining values in variables, you can easily change the values in one place without going through your entire code to update them.

Data Structures

- R offers a variety of data structures to handle and manage data.
- The structure you choose depends on the type of data and how you intend to use it.
- Common structures include vectors, matrices, data frames and lists

Data Structures

- Vectors are the simplest and most common data structure in R. They are one-dimensional and can only contain elements of the same data type.
- Matrices are two-dimensional structures with rows and columns. All elements in a matrix must be of the same type.
- Data frames are two-dimensional structures like matrices but can contain different types of data in each column.
- Lists are the most general structure. They can contain elements of different types (e.g. numbers, strings, vectors and matrices, and other lists)

First Plots

- R offers tools for data visualization using base R or external libraries like ggplot2.
- Base R functions can create many types of plots such as scatter plots, line plots, and histograms.
- The core function for plotting in base R is plot() .

First Plots

- Scatter plots (observations are displayed as points) are the default when using the plot() function.
- The plot type can be changed using the `type` argument (e.g. `type="l"`)
- Titles, labels, and axis limits can be changed using additional parameters `main`, `xlab/ylab` and `xlim/ylim`, respectively.
- Line types, point characters and colors can also be adapted by using the optional arguments `lty`, `pch` and `col`, respectively.
- In the tutorials you'll also see how to add multiple lines to one plot, add legends or display multiple plots side by side.