# R Programming:

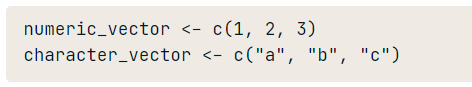
R is an open-source programming language and environment for data analysis, data visualization, machine learning, and statistics. Widely used for developing statistical software and performing data analytics, it is especially known for its ability to create compelling visualizations, giving it an edge over some of the other languages in this space. Some of the key benefits of R include the following:

* It is an open-source platform-independent programming language,
* It can be paired with many programming languages, including Python,
* It is highly extensible, which means developers can continue to add functionalities by defining new functions,
* It facilitates the handling of structured as well as unstructured data which means it has a more comprehensive data capability,
* It has libraries such as Ggplot2 and Plotly that offer aesthetic graphical plots to its users, we can make reports with the data and scripts embedded in them; also, interactive web apps that allow users to play with the results and the data,
* It is dominant among other programming languages for developing statistical tools.

## Vectors:

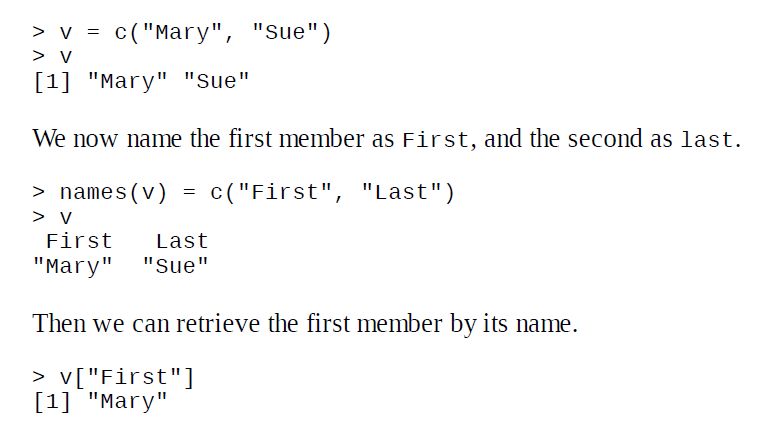
Vectors are one-dimension arrays that can hold numeric data, character data, or logical data. In other words, a vector is a simple tool to store data. For example, we can store your daily gains and losses in the casinos.

In R, we create a vector with the combine function c(). We place the vector elements separated by a comma between the parentheses. For example:



## Named Vector Members:

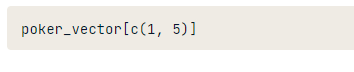
We can assign names to vector members. For example, the following variable v is a character string vector with two members.



Vector selection: the good times

**How about analyzing your midweek results?**

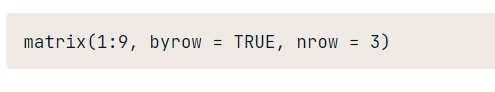
To select multiple elements from a vector, you can add square brackets at the end of it. You can indicate between the brackets what elements should be selected. For example: suppose you want to select the first and the fifth day of the week: use the vector c(1, 5) between the square brackets. For example, the code below selects the first and fifth element of poker\_vector:



## Matrix:

In R, a matrix is a collection of elements of the same data type (numeric, character, or logical) arranged into a fixed number of rows and columns. Since you are only working with rows and columns, a matrix is called two-dimensional.

You can construct a matrix in R with the matrix() function. Consider the following example:



In the matrix() function:

The first argument is the collection of elements that R will arrange into the rows and columns of the matrix. Here, we use 1:9 which is a shortcut for c(1, 2, 3, 4, 5, 6, 7, 8, 9).

The argument byrow indicates that the matrix is filled by the rows. If we want the matrix to be filled by the columns, we just place byrow = TRUE.

The third argument nrow indicates that the matrix should have three rows.

**Q. Construct a matrix with 3 rows containing the numbers 1 up to 9, filled row-wise.**

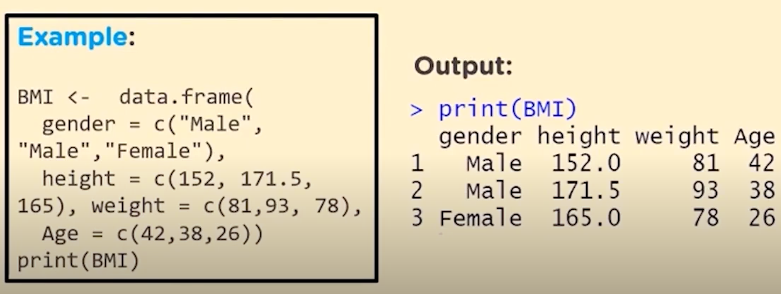
****

**List:**

* A list is a generic vector that can contain object of different types.
* We use the list() function to create a data frame.

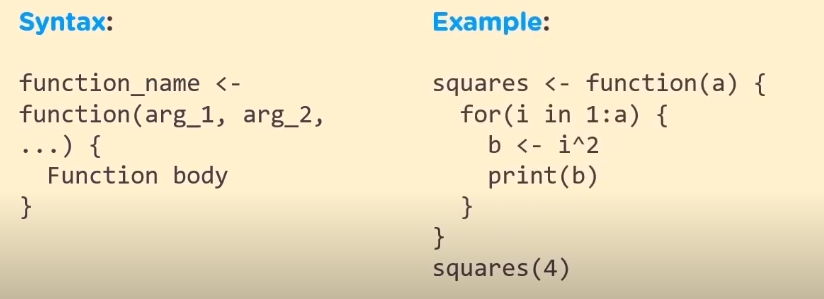
Data frame:

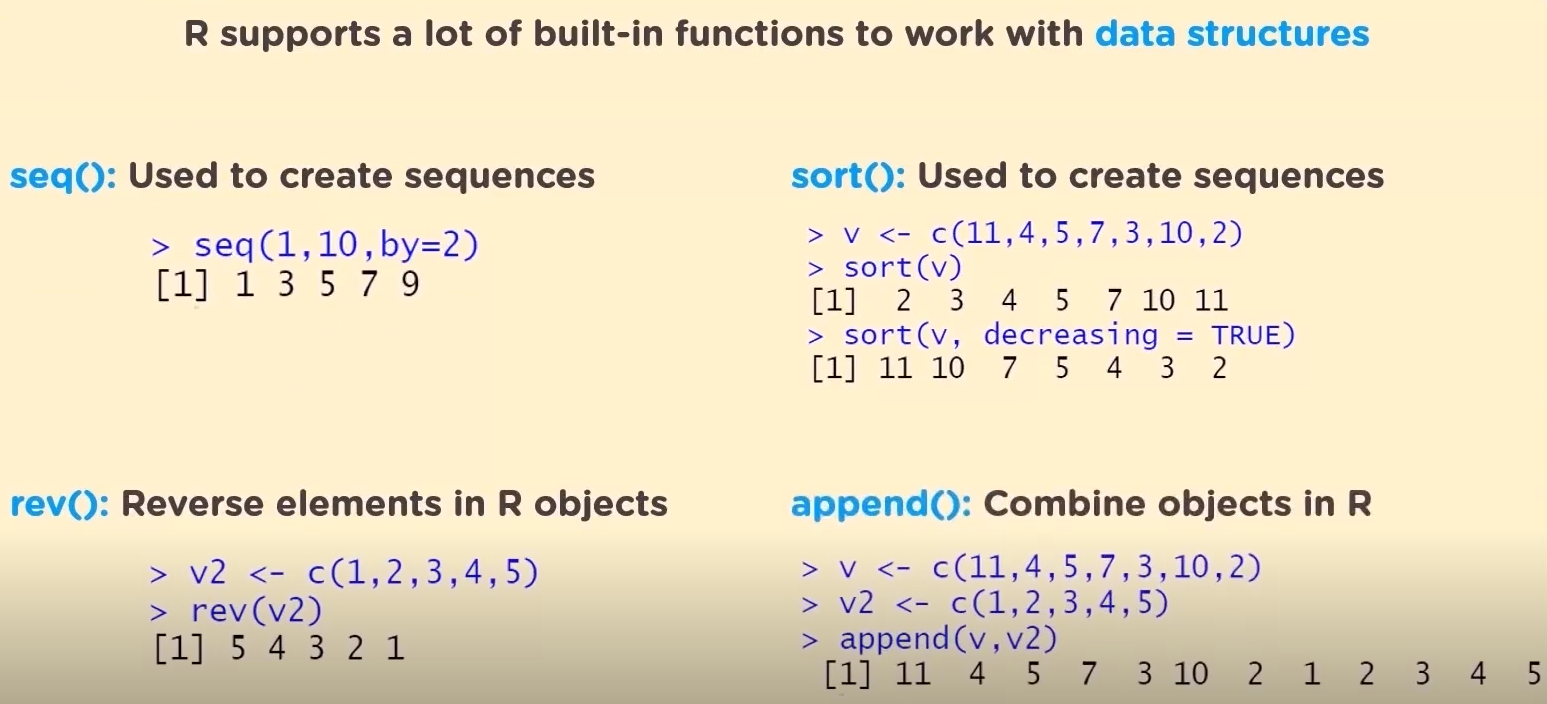
* A data frame is used to store the data in the form of table
* We use the data.frame() function to create a data frame



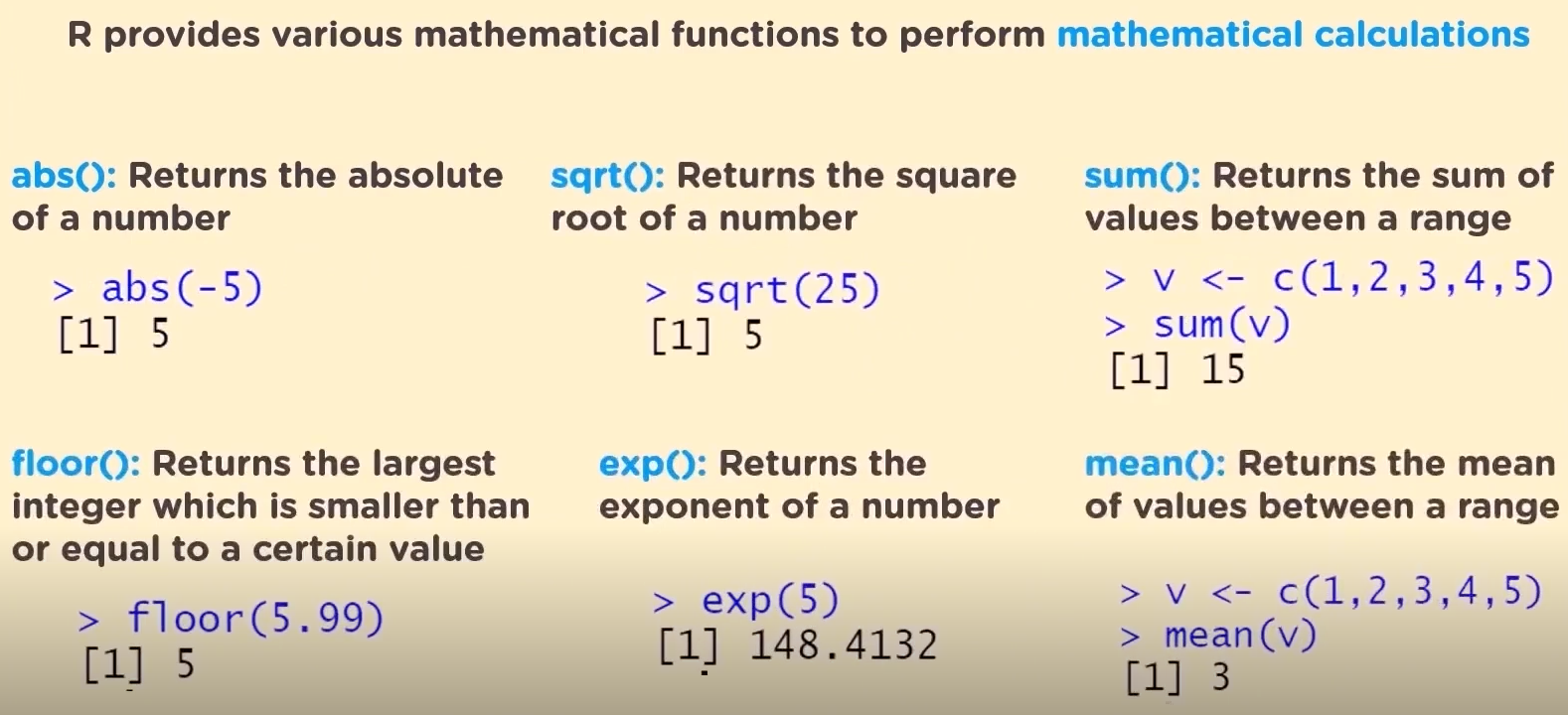
## Functions:

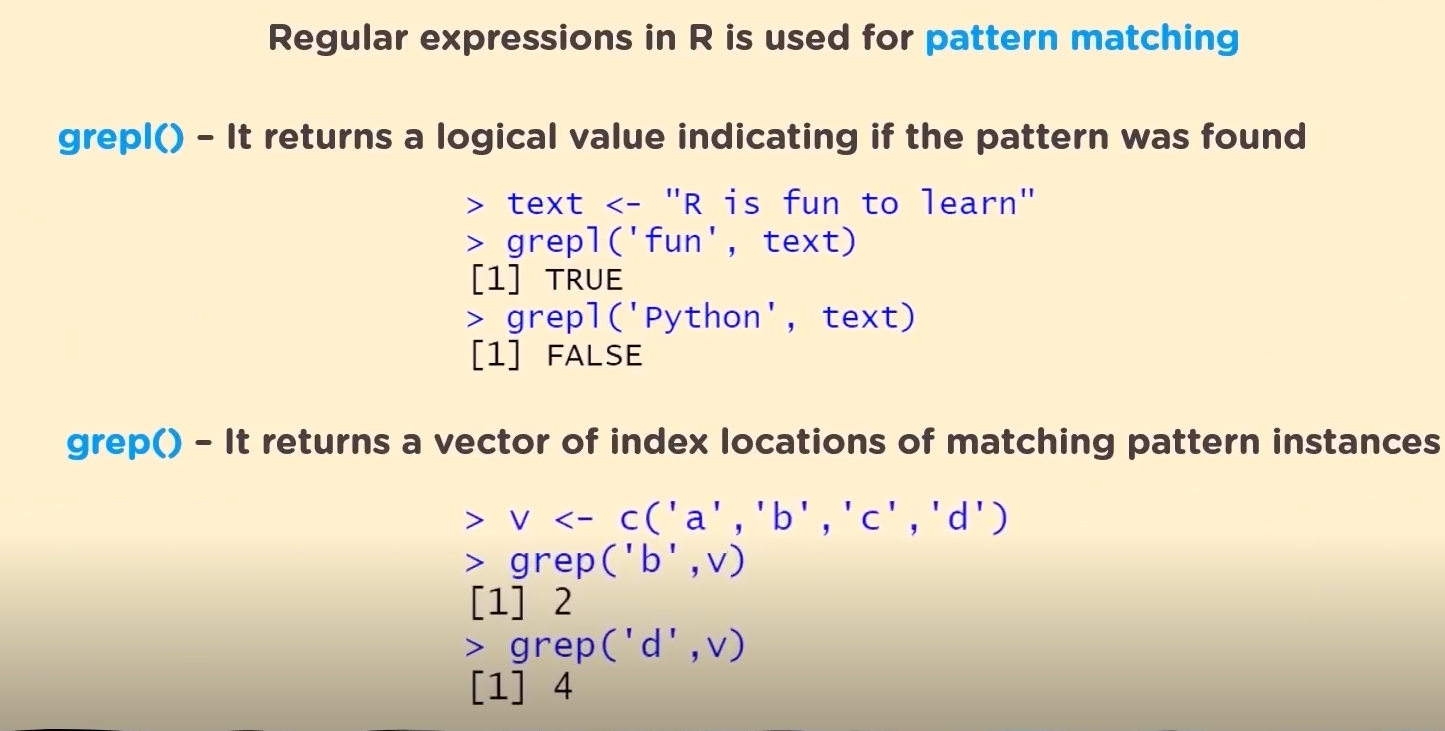
A function is a set of statements to perform a specific task. R has a large number of in-built functions and the user can create their own functions.





Some mathematical functions for data science:





# Data Manipulation in R – dplyr

The dplyr package is used to transform and summarize tabular data with rows and columns.

