R for beginner course

BY

KASONGI, NGWINAMILA

DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL STUDIES, UDOM

Overview:

Today's sessions we shall cover the followings:

1. Explore user interface of basic R and graphical interface for R called R studio
2. Explore some basic concepts that you must know to get started with coding in R

Note: learning coding requires a lot of commitment and perseverance

Have created a googled rive folder where you can access R codes, presentations and data

Use this link: xxxxx

Types of variables

Integer/numeric - whole number e.g., 2 , 3, 5 etc.

Floats – decimal number 1.2, 2.5 etc.

Character – strings or text e.g., ‘geography’ or “geography”

Boolean – TRUE or FALSE

Date – 02/03/2003

Factor – categorical variable e.g., Sex (Male, Female)

Comparison operators

> greater than

< less than

>= greater or equal

<= less or equal

!= not equal

== equal

We use comparison operator to compare two items/variables

E.g., 20 > 10, 50 != 51

Boolean operators

AND = &

OR = |

Not = Not

Usually, we use Boolean operators to concatenate two or more strings/character variables

We can use vein diagram to understand the Boolean operators

E.g., ‘Ugali’ OR ‘Rice’

Data objects

We use data objects to store data (variables) such as numbers, strings etc.

Common data types in R:

- Vector
- List
- Matrix
- Arrays
- Dataframe

Vector

It contains similar atomic datatype (variables)

E.g., names

c(‘R’, ‘SPSS’, ‘STATA’)

E.g., numbers

c(1,2,3,4,5)

List

It contains elements of different types like − numbers, strings, vectors and another list inside it

E.g., list(1,2, ‘geography’, ‘r course’)

Matrix

The elements are arranged in a two-dimensional rectangular layout (row and column).

They contain elements of the same atomic datatype types.

Though we can create a matrix containing only characters or only logical values, they are not of much use

We use matrix()

1	2
3	4

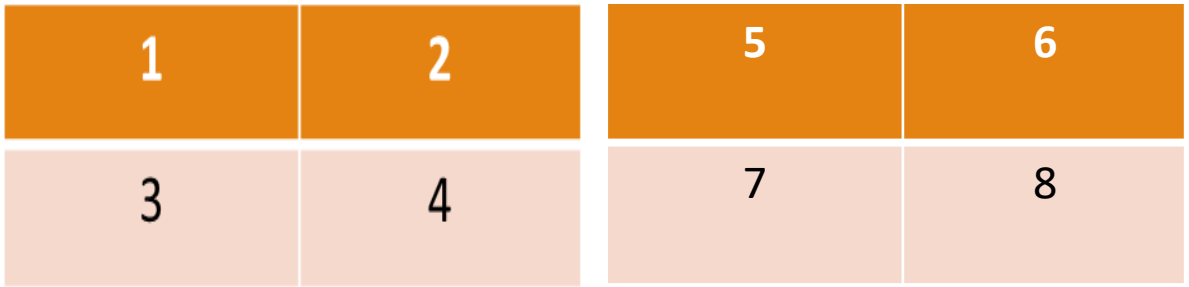
Arrays

The elements are arranged in more than two-dimensional rectangular layout (row and column).

They contain elements of the same atomic datatype types.

It is important to specify number of dimensions when creating array object

We use array() to create an array object



Dataframe

It’s like table which has rows and columns.

Each column store specific type of variables

It can take different atomic datatypes(variables)such string, number etc.

Most of our dataset are in dataframe object

We use data.frame() to create dataframe object in R

Sex	Age
Male	20
Female	45
Female	18

If statement

We use if condition to execute line of codes when conditions are met

E.g., If temperature is greater than 30, say it’s hot day

If(condition){execute if condition is true}

If(condition){execute if condition is true} else{execute if condition is not true}

Iterations or loop

Iteration refers to repeatedly execution of code.

A common loop in R is a “for loop”

Loops helps to iterate over same operations without writing too many codes – code simplicity

for(i in x) {execute the code }

Let create a for loop that multiplies 2 to each item in our vector

Functions

It sequence of program instructions that takes a user input(s) and returns certain output(s)

We use a lot of built-in functions for various purposes

However, there are times you need to create your own function e.g., for iteration purposes

function(inputs){execute the code}