R Introduction

Young W. Lim

2024-06-24 Mon

Outline

Based on

"An Introduction to R" Notes on R: A Programming Environment for Data Analysis and Graphics
W. N. Venables, D. M. Smith, and the R Core Team

- the <u>entities</u> that R <u>creates</u> and <u>manipulates</u> are known as objects
  - variables
  - arrays of numbers
  - character strings
  - functions
  - more general structures built from such components
- during an R session, objects are <u>created</u> and <u>stored</u> <u>by name</u>

`https://www.w3schools.com/statistics/statistics_statistical_inference.php`

- The R command
  ```
  > objects()
  ```
  (alternatively, `ls()`) can be used
  to display the names of (most of) the objects
  which are currently stored within R.

- the <u>collection</u> of objects currently stored
  is called the workspace

`https://www.w3schools.com/statistics/statistics_statistical_inference.php`

Types of R objects (1-1)

- a vector is an ordered collection of numerical, character, complex or logical objects.

  vectors are collection of <u>atomic component</u> or modes the same data type

- a matrix is a multidimensional collection of data entries of the same type.

  matrices have two dimensions.

  <u>rownames</u> and <u>colnames</u>

`https://www.w3schools.com/statistics/statistics_statistical_inference.php`

- a list is an ordered collection of objects
  that can be of different modes
  different data types

- though a data.frame is
  a restricted list with class data.frame,
  it maybe regarding as a matrix with columns
  that can be of different modes.
  It is displayed in matrix form, rows by columns.
  (Its like an excel spreadsheet)

- A data.frame is a list of variables of the same number of rows
  with unique row names, given class data.frame
  if no variables are included,
  the row names determine the number of rows.

https://www.w3schools.com/statistics/statistics_statistical_inference.php

Types of R objects (2)

- A factor is a vector of categorical variables,
  it can be ordered or unordered.

- array an array in R can have one, two or more dimensions.
  useful to store multiple related data.frame
  (for example when I jack-knife or permute data).
  Note if there are insufficient objects to fill the array,
  R recycles (see below)

https://www.w3schools.com/statistics/statistics_statistical_inference.php

Dataframe and class objects

- By definition, a class is a code template for creating objects.
- This means that you can define a class
  that will create a certain object for you
  when this class has been instantiated.
- Then, the DataFrame is a type of pandas object.
- Therefore, you can say there's the pandas DataFrame class,
  that is code template that can create a DataFrame for you.
  - pandas is a fast, powerful, flexible and easy to use open source data
    analysis and manipulation tool, built on top of the Python
    programming language.

`https://365datascience.com/question/difference-between-dataframe-and-class-objects`

1. Basic Attributes
   - The most basic and fundamental properties of every objects is its mode and length
   - these are intrinsic attributes of every object.
     Examples of mode are "logical", "numeric", "character", "list", "expression", "name/symbol" and "function".

https://www.w3schools.com/statistics/statistics_statistical_inference.php

Attributes of R objects (2)

1. Basic Attributes (continued)
   - character: a character string
   - numeric: a real number, which can be an integer or a double
   - integer: an integer
   - logical: a logical (true/false) value

`https://www.w3schools.com/statistics/statistics_statistical_inference.php`

1. Basic Attributes
2. Other Attributes, dimension

Object	Modes
vector	numeric, character, complex or logical
matrix	numeric, character, complex or logical
list	numeric, character, complex, logical, function, expression, . . .
data frame	numeric, character, complex or logical
factor	numeric or character
array	numeric, character, complex or logical

- Whether object allows elements of different modes. For example all elements in a vector or array have to be

of the same mode. Whereas a list can contain any type of object including a list.

https://www.w3schools.com/statistics/statistics_statistical_inference.php