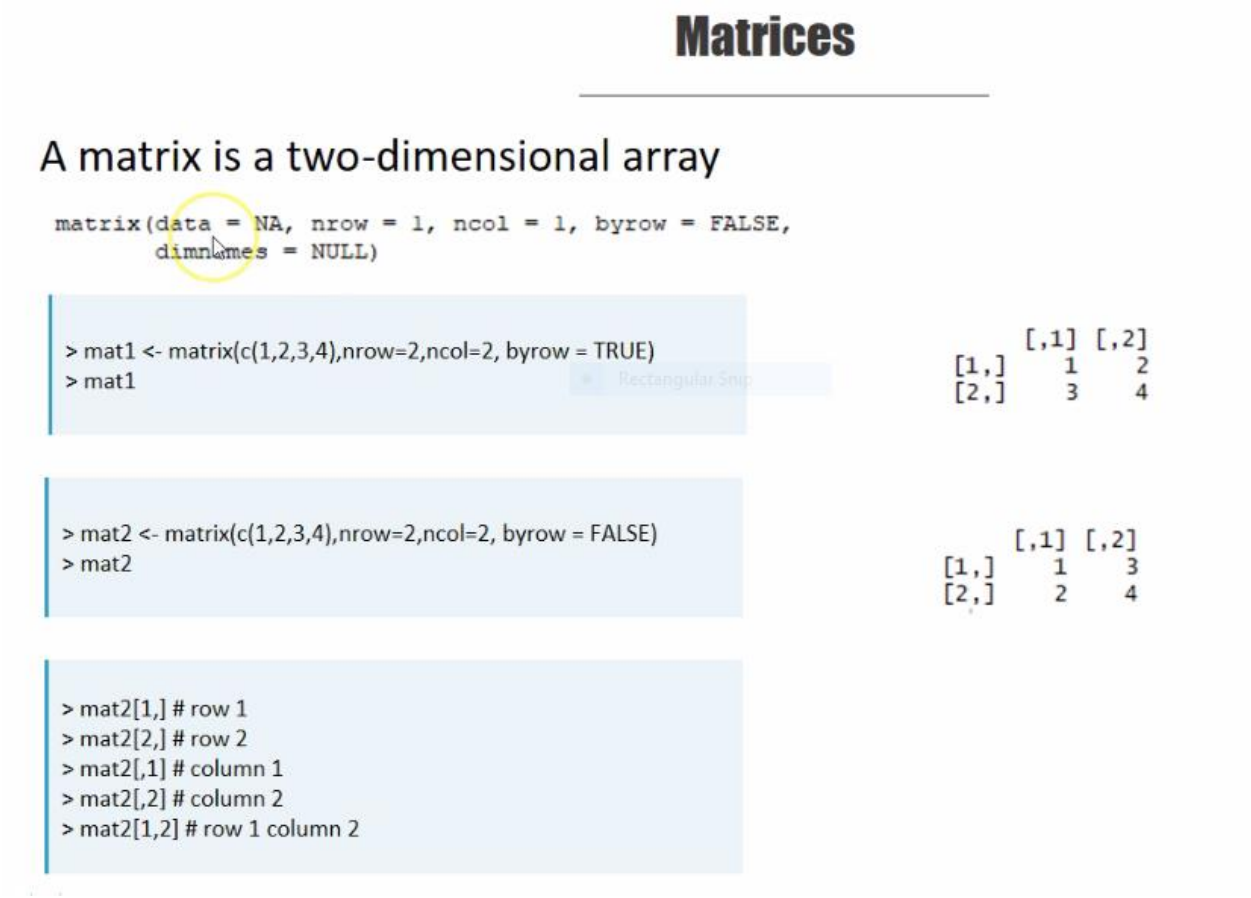
Matrix, Array and Data Frame - Study Note

Matrices:

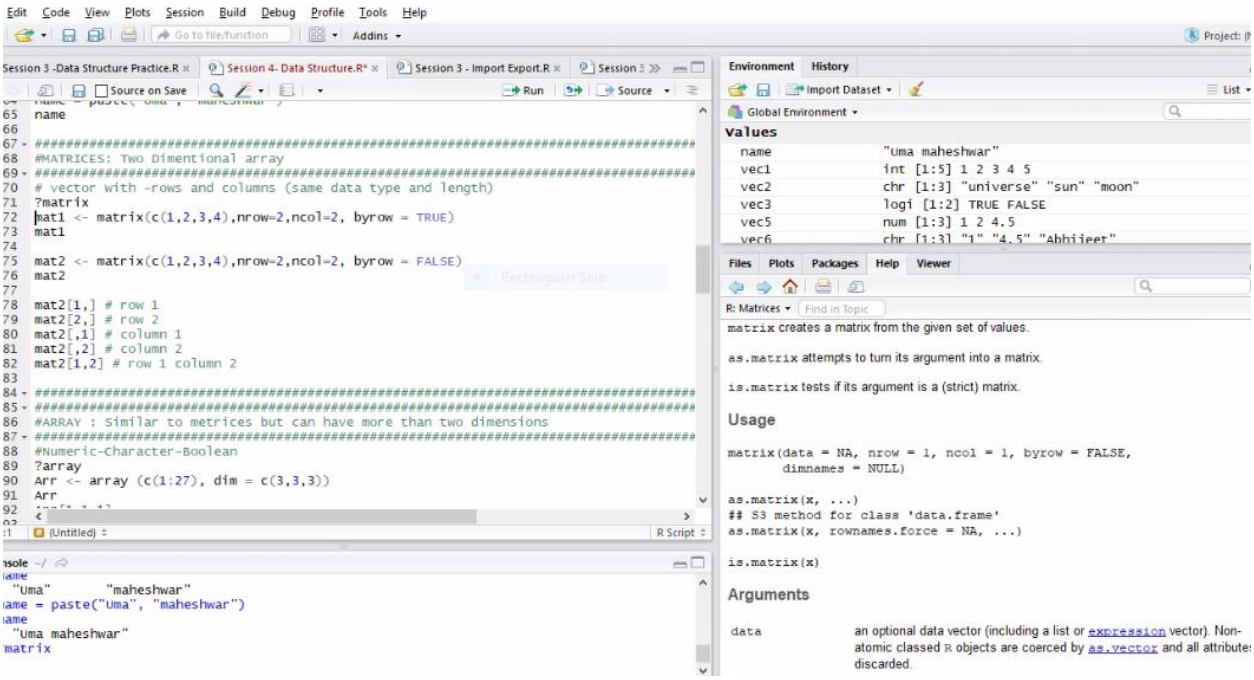
Matrices are the R object, which is a collection of data elements arranged in a two-dimensional data array.Although we can create a matrix containing only characters or only logic values which are not of much use. We use matrices containing numeric elements to be used in mathematical calculations.

Matrix is created using the matrix() function.Matrix() function is being used to create a matrix. Here we show an argument.

Here,the argument is matrix(data=NA,nrow=1,ncol=1,byrow=FALSE,dimnames=NULL). We have a data, word data which we want to create a matrix, suppose 1-4 we want to create a matrix. Then, how many rows here, there are by default 1 row but we are changing it two and by default it’s column is 1 but we are also changing this to two and suppose by default byrow is equal to False and by row is nothing but suppose 1-4 you want to create a matrix then how you want to do it? Either this way byrow equal to TRUE equal to 1,2,3 and 4.



But if you change that byrow equal to FALSE then you want to do it by column so first 1 then 2 then 3 then 4 that's how you can do it. Here, we are keeping all the numerical values but here we are showing it 2 rows and 2 columns and by rows equal to TRUE that means we will start from the row and then in the second row. Similarly, if you access any element of a matrix.

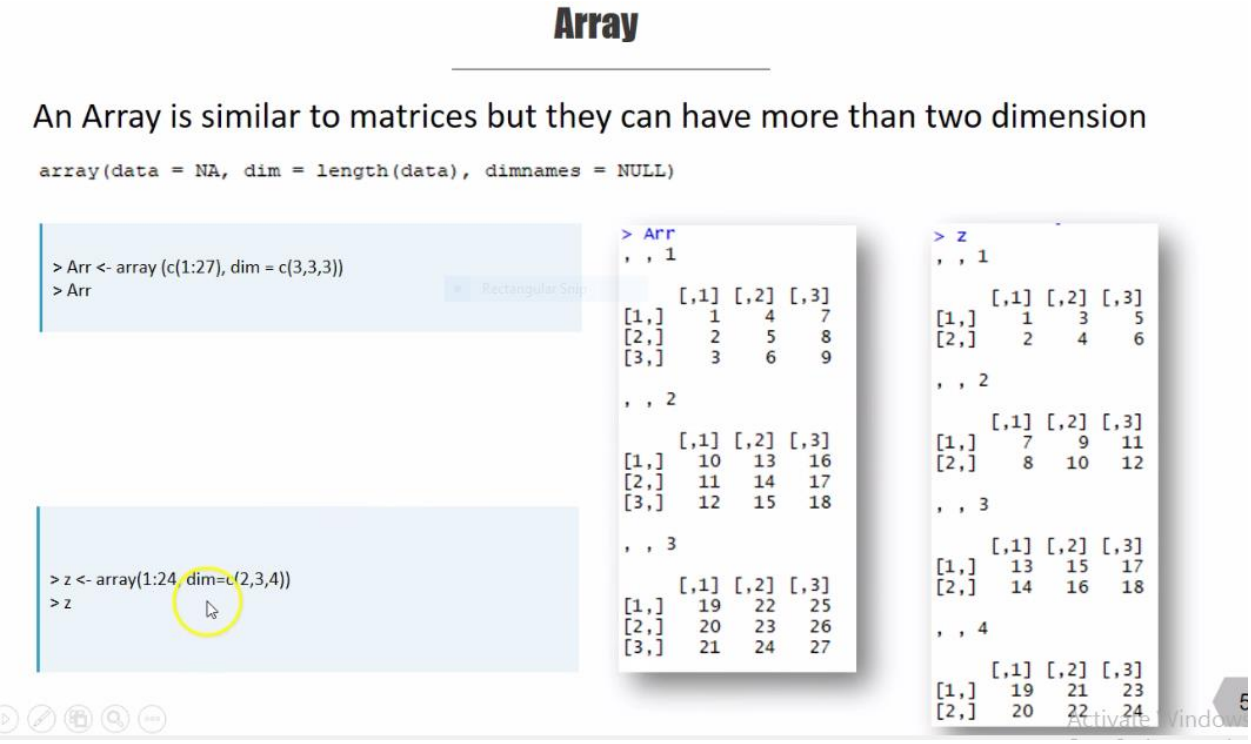


Suppose, we are creating matrix 2 where we access the first row then you can say 1,so this is row, column so we are asking for the first row and it would give us first row, the second row then it also gave us first column because we said row, column then second column.

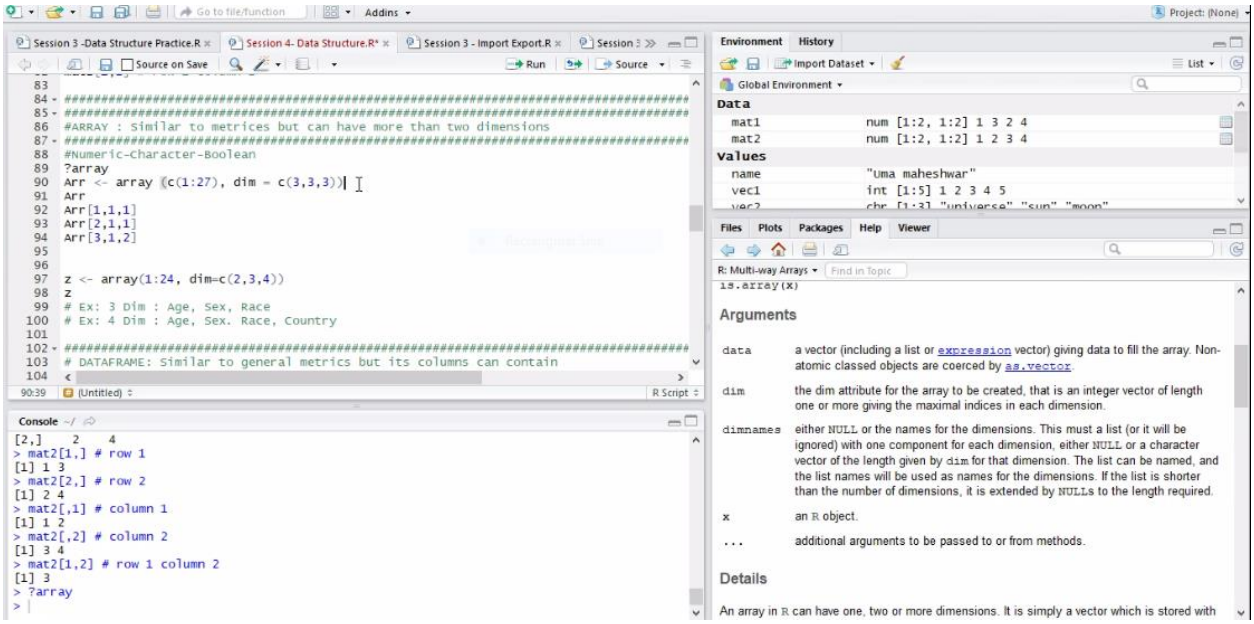
We haven't written anything in a row that's why it would give us all the rows and only it would give us first column. Similarly, it would give us all the rows but an only second column.

Array:

Now we have discussed the Array. An Array is similar to matrices but it can have more than two dimensions. Here, you can store 2*3*4 anything you can create. R Array is the data objects which can store data in more than two dimensions. An Array is created using the Array() function. The array can store only data type. Array takes vectors as input and uses the values in the dim parameter to create an Array.



It can contain multidimensional rectangular shaped data storage structure. “Rectangular” in the word, each row is having the same length similarly for each column and other dimensions. But Array can store only values which have similar kind of data,i.e. variables/elements having a similar data type.We create Array using the Array() function. The argument is array (data=NA, dim=Length(data), dimnames=NULL).

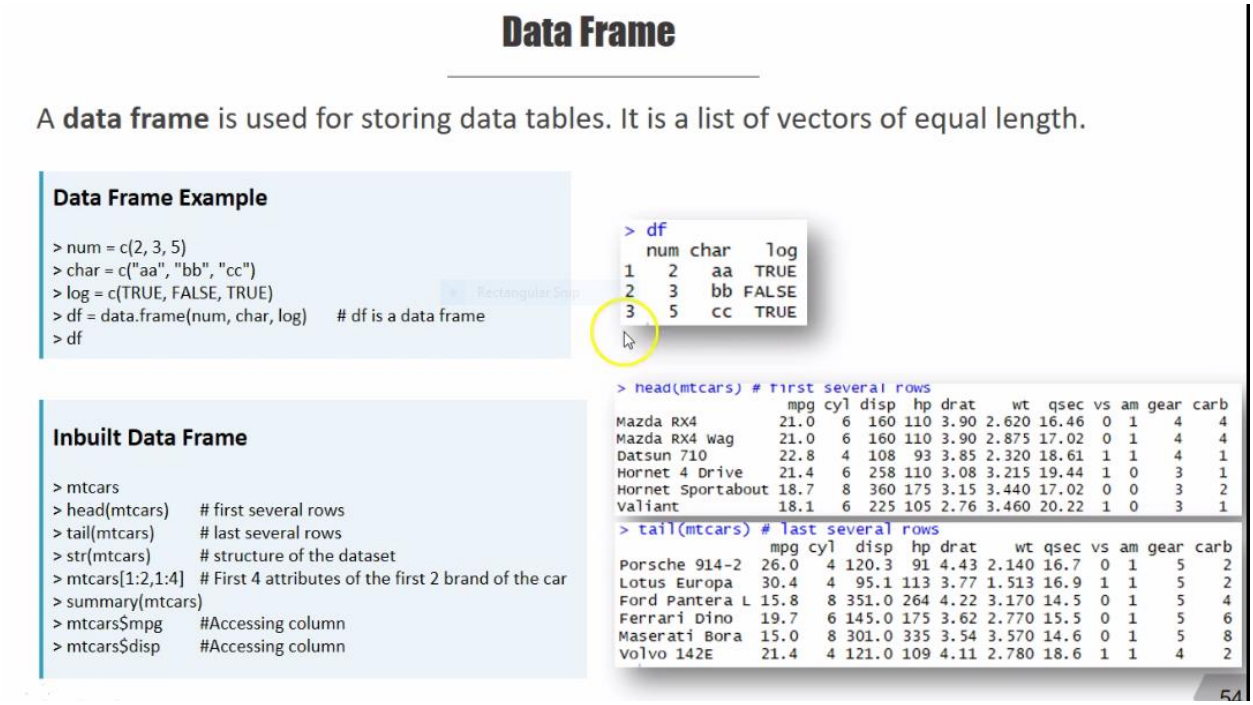


We first pass the data here we pass 1 to 27 with this dataset we want to create the Array. What is those dimension? The dimension is 3*3*3 here you can store salary, age, designation, grade or something like that. It is pretty simple and you can see the result as well.

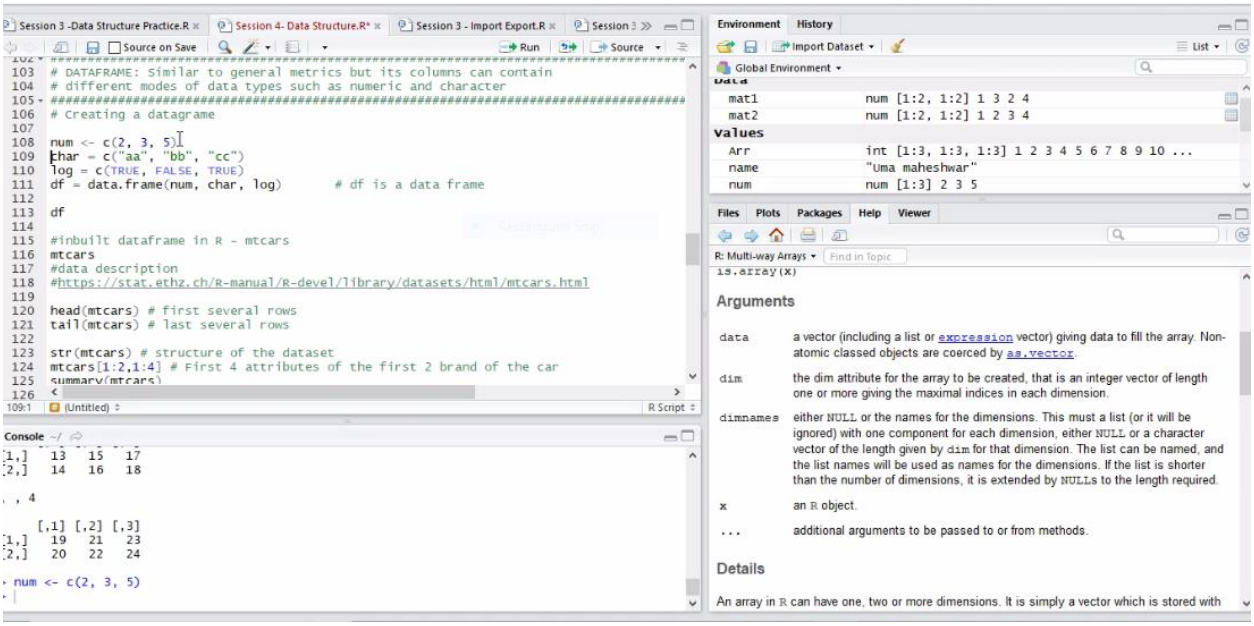
Dataframe:

Dataframe is a table or two-dimensional array like structure where each column contains values of one variable and each row contains one set of values from each column.

Data frame is used for storing a data tables. This is a list of vectors of equal length. All the vectors we have learned previously like Numerical vectors, Character Vectors, Logical Vectors and all those things. Now you just mix them the three vectors you mix them. Suppose you want to create a data frame where you have one attribute as Numerical like age or salary which is numerical but you have three employees so all of this vector like the same length. So you have three employees then you want to store their gender as well male, female. Similarly, you want to understand whether they left the company or not something like that TRUE FALSE or something like that. If you achieve this kind of data set then you have to use data frame.



Now we use data. frame, we use this data. frame function. Before that we have to three vectors first, suppose we are storing 2,3,5 a vector called num then “aa”,”bb”,”cc” in char and then TRUE, FALSE and TRUE in logical then ultimately using data.frame()function and creating that data frame. Here we are using the data frame name df. If you want to see what is df? Or what is being stored in df? Thereafter you get the result.



Think about a business scenario, where you have a lot of employees and here you use the data frame. That's why data frame is one of the most important structures in R. In R we have a lot of inbuilt data frame which you can explore. One of the popular dataframe is mtcars. Mtcars is a dataset of a car distribution. We have df which is already been created in R that's why we said this inbuilt database. Here mtcars has 32 model of cars. If you want to see a couple of them then you use head() function, you can use the head() function for first several rows. similarly, you can use tail() function to see last couple rows. The tail function gives you the below of the lists. Another function is str() function which is the structure of a dataset. The str() function gives you the detail observation. For str() function you get a sense how your data set looks like. Likewise, you can see the summary which is better than the full data structure. The summary() function gives you the min or max from each of those attributes.

This brings an end to this post, I encourage you to re read the post to understand it completely if you haven't and THANK YOU.