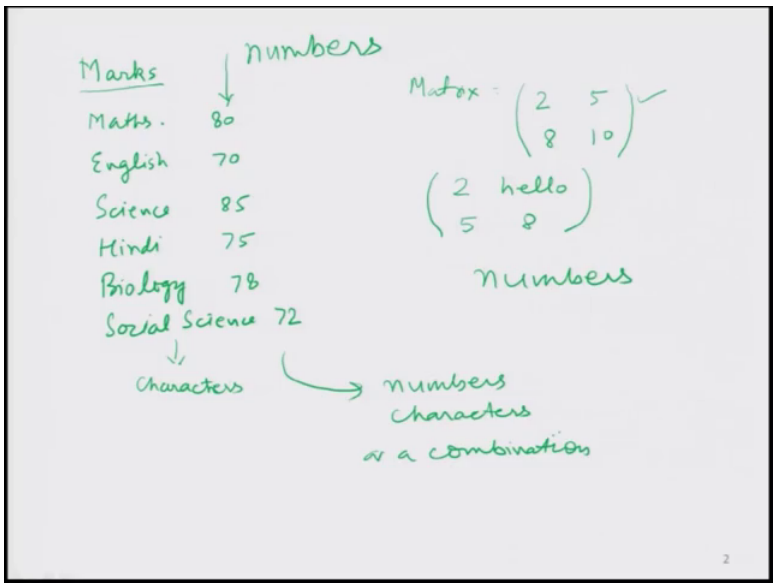
Introduction to R Software
Prof. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture - 18
Lists

Welcome to the next lecture on the Introduction to R Software. In this session we are going to discuss about a new concept that is called as a List. First let us try to understand what do we really mean by list, and how it is different from say some other commands like as vectors or matrix. And this topic will be covered in the 2 lecture this lecture and the next lecture.

So, first let us try to understand what is called a list. Let me take a very simple example to understand what do we mean by here list. Suppose you are asked to write down the marks that you have obtained in say class 10.

(Refer Slide Time: 01:09)



What you will do, here you will try to write down the subjects say Maths, then say some language, say English, say Science, say Hindi, Biology, and say Social Science and so on. And then you will write the marks obtained in each of this subject say out of 100, suppose the candidate has got 80 out of 100, in English he is got say 70 out of 100, in science say this 85 marks out of 100, in Hindi suppose 75 marks out of 100, biology say this 78 marks out of 100, and say social science say 72 out of 100.

Now, in this a structure what do you find, that I have here 2 types of structure: one these are here some characters, and on this side we have here numbers. Now so, if you try to see we have written a combination of characters and numbers, and what do you say in simple language, I have created a list of the marks in every subject which a candidate has got in class 10. This is simply a list. Now you will get is a little bit puzzled, what is the difference between this lists and say any other thing.

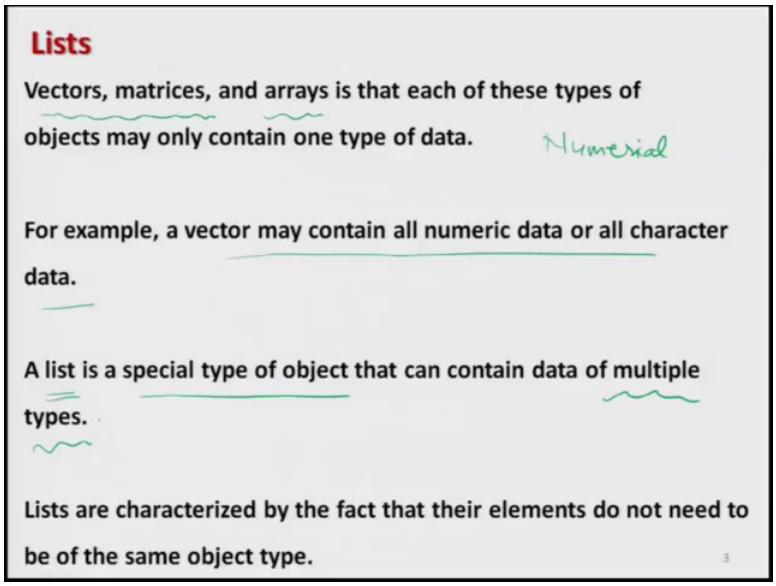
So, let me try to take here an example of here matrix; matrix, is simply an arrangement of numbers in rows and columns for example, this number can be 2 5 8 10, and all sorts of mathematical manipulations can be operated over this matrix, here is in case if I try to write down here a matrix something like 2, and say here I write hello, and say here 5, and say here 8, then will you call this is a matrix. Well, one can call, but you cannot operate any mathematical operators over this matrix.

So, in say matrix all the entries are only in numbers, and when I am trying to deal with list can have the entries in numbers as a character, or a combination of them right. So, whenever we are trying to do the data analysis, there are certain situations where we need to input the data, or we need the output of data, in the form of a list.

So, the advantage of having a list is that, we can have all sorts of data inside the list that can be a character that can be a say number or say anything else, whereas when we are trying to deal with vectors or say matrix, then we can enter the data usually in the form of only numbers, that can be used with any mathematical operators.

So, now first let us try to have a brief discussion, and say idea about the list, and then we will see how to create a list and how to manipulate over the list in say r programming right.

(Refer Slide Time: 05:43)



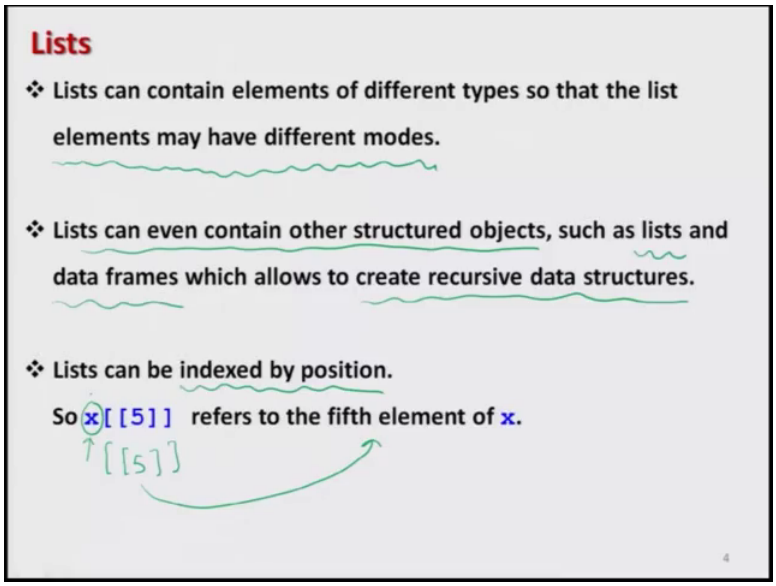
So, now, as we have discussed here we have understood that, whenever we are calling of the vectors matrix or say array usually in this type of objects, that data can only be of one type, and usually it is numerical whenever we think about matrix or vector usually we think in terms of some numbers.

On the other hand in case if you want to use a vector or a matrix that can also contain all the characters also, in case if you need, but as we said this will not be useful from the mathematical point of view for example, in case of a matrix has a character say hello you cannot obtain say x transpose x or you cannot multiply that matrix with anything else.

So, in order to overcome such issues list is used, this list is a special type of object that can have multiple types of data, or the data in multiple modes. What is called a mode, we already had discussed the concept of mode, when we were trying to discuss about the logical vector, but in the but after few slides I will try to take a this issue in more detail.

So, one of the basic characteristic that you have to keep in mind about list that in the list the elements do not need to be of the same type, they can be of different types. And that is the advantage of having a list over a vector matrix or a array.

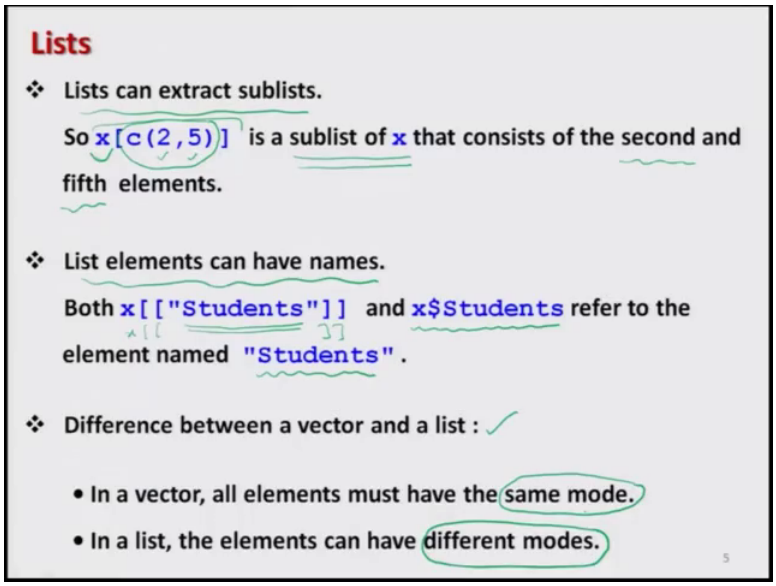
(Refer Slide Time: 07:37)



And when we are saying that list can contain different types of data; that means, in a technical language we are saying, that the elements of the list may have different modes the advantage of having a list is that once we have a list, we can use the elements inside the list to create another structures of data set, and even list can also contain some structure objects, that itself can be a for example a list matrix data frames and so on, and using this bigger list we can create different types of data structure.

The list also has an index option; that means, the entries in the list can be indexed by their positions for example, this was the sake of illustration, if I say here if I try to create a list here, say here, denoted by here, is x then, if I try to write down inside this square brackets w square brackets and if I try to write down here 5, then this 5 is referring to the 5th element in the variable x. So, that is another advantage of using a list that all the entries can be indexed.

(Refer Slide Time: 09:24)

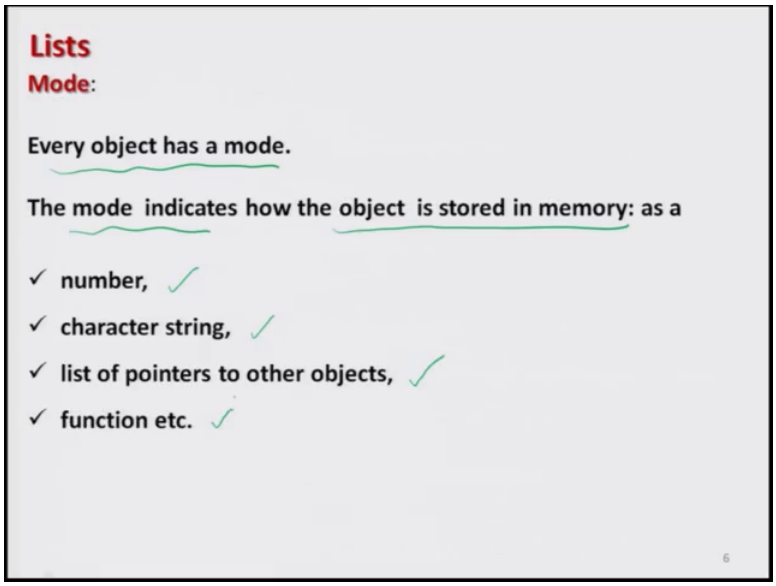


Lists has another advantage that a list can extract sub list; that means, list can also extract a part of the list that we can call is as say sub list. Suppose there is a list which is given by the name x. And suppose I want to extract its 2nd and 5th element.

So, this is a sort of here sub list of x, and this sub list of x can be extracted using the kiss command. So, this is a sub list of x, that contains 2nd second and 5th elements of x, another advantage of using lists is that the elements of list can also have some names for example, if I try to give a name here students, then these names are enclosed by the double quotes, and they are enclosed in the double square brackets sign over here, followed by the name of the variable that is containing the list. And this will refer as if I am trying to have the name of students, and that we had seen earlier. And we will see in the further lectures also that this thing can also be obtained using this command x dollar students. And this means that I am trying to refer to an element whose name is students.

So, the basic difference between a vector and the list is that in a vector, all the elements are going to have the same mode, and in a list the elements can have different modes. So, now before going further let us try to understand, what do we really understand by modes.

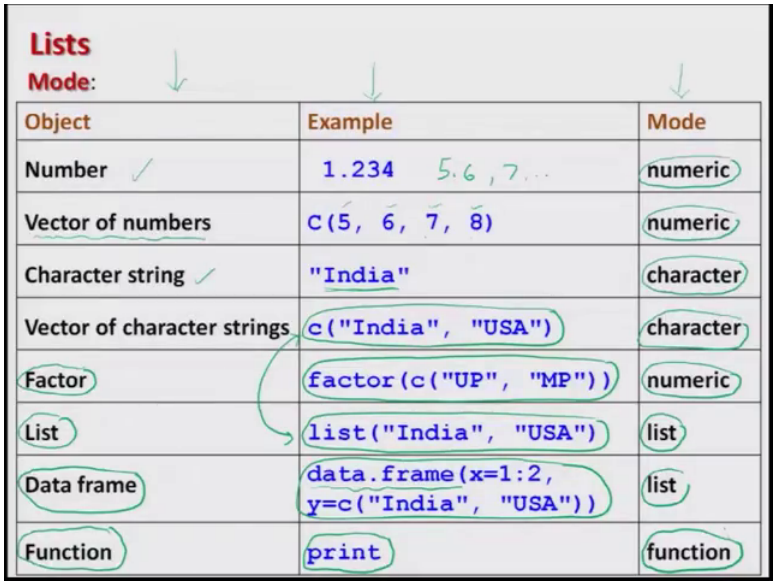
(Refer Slide Time: 11:42)



Actually every object has a mode, what does this mean; the mode indicates that how the object is stored in the memory of a computer.

For example this can be stored as a number, as a character string, as a list of pointers to other objects, as a function, or say something else also. So, whenever we are trying to use the function, let us try to understand, what are the different types of modes?

(Refer Slide Time: 12:23)



So, here in this table I have listed the possible modes, and then I will try to take an example that what do we mean by that particular type of mode and, what is the mode that

is mentioned in the 3rd column. So, for example, an object can be a number for example, 1.234 5.6 or say 7, and so on. What is this that, we know from our common language that this is a number, so the mode of this number is called as numeric.

Similarly in case if I try to take a vector of numbers; that means, I am trying to contain I am trying to consider more than one numbers at a time for example, I have written here a vector which is combined by the 4 values 5 6 7 and 8, so you can see here 5 6 7 and 8 all are actually numbers so, a combination of a number is also number, so the mode of vector of numbers is also numeric.

Now, let me try to take some characters, which are non numbers for example, if I try to take a character string say some name say India then; obviously, this is not a number, but this is a character. So, I can say the mode of our character is thing is character, and similarly if I try to take more than one characters at a time, they can be combined through the c command in a vector and for example, here I try to take here 2 characters India and USA and I try to club them, combine them, with a character c at the operator c, so this is the combination of the characters. So, the combination of a character is also a character. And hence the mode of vector of characteristic is a character. There is another type of object what we call as factor, we have not yet discuss about the factor that we will take up in the upcoming lectures.

But here in case if I try to take a factor or a combination of factor, this can be written in this format, and you have to just note down that the mode of a factor is numeric. Well, at this stage it is difficult for me to give you more details, but definitely as soon as we try to discuss the concept of factor I will try to take up this topic again. So, here you could just try to keep in mind.

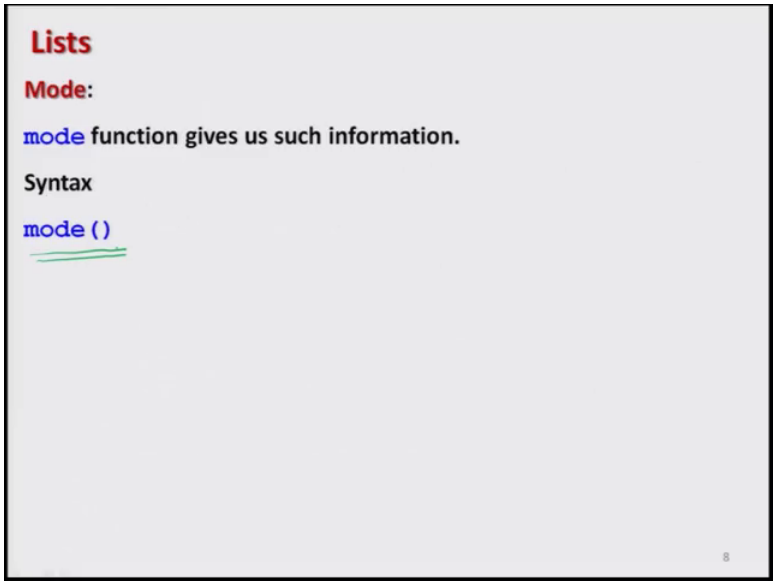
Now, I come to another this object, which is here list. List can contain all sorts of element, they can be character strings they can be numbers or say anything for example, I am trying to create here a list of India and USA and, the mode of list is actually list, at this moment you may get little bit confused, what is the difference between this and this they are trying to give us the same thing, but you just wait for some more slides and he will come to the discussion on list, and then it will be more clear to you that what is the difference between this. What I am trying to tell you here that whenever we are trying to create a list the mode of list is list.

And, similarly there is another object what we call as data frame, we have not done data frame up to now, but we will be doing it in the upcoming lectures, forth coming lectures. So, till then you have to wait to understand what do we really mean by data frame, but anyway means in a data frame we can also contain different types of data, that can be character string that can be numerical and they are denoted by data dot frame over here, and the mode of data frame is also a list.

Another object is function, we already have done this function in initial lectures and we had used one function there as say here print, that was used to display the output or to display anything any number, or a character and if you try to find out the mode of this function, then this is a function, and these type of modes help us in determining in a bigger data set, where we cannot look into the individual value to find whether my data set has a number or a character, or a combination of them, or it is a logical vector and so on.

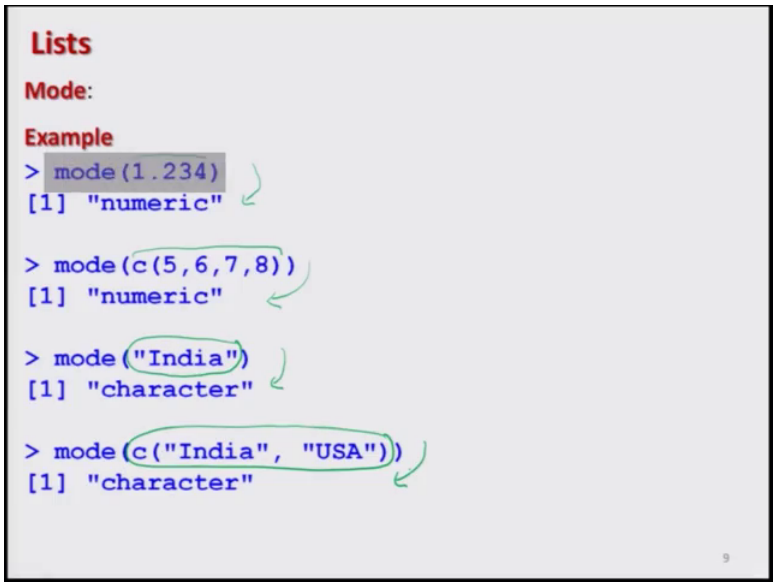
So, I have taken some of the objects some of the important objects, and then I have described you that what are their modes right, now we come to the aspect that how do you find out the mode, well we also had discussed it earlier, but here I will try to repeat it with some more example on other types of mode.

(Refer Slide Time: 18:43)



So, mode function gives us the information about the mode, and the syntax is simply here mode and, inside the bracket you have to a specify the value of which you need to find out the mode.

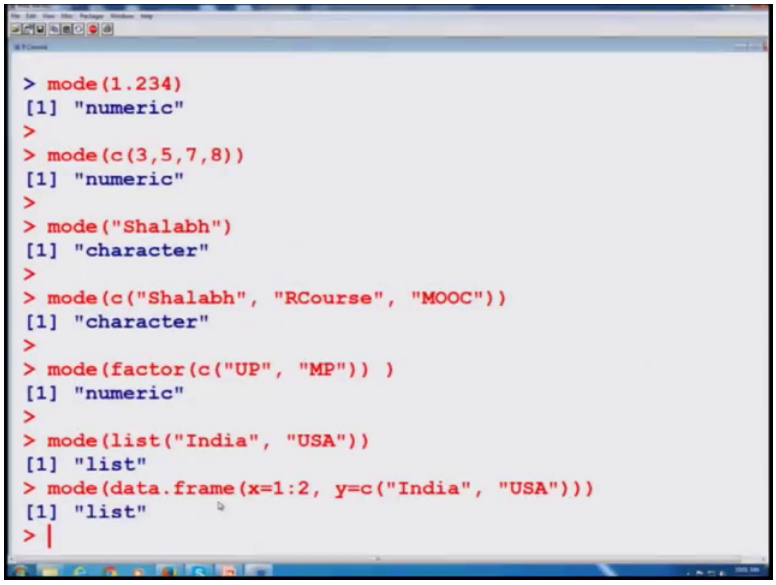
(Refer Slide Time: 19:03)



So, let me try to take here some examples and we try to see here for example, I am trying to take the same example that I have given in the table, for example, can I want to find out what is the mode of 1.234 as soon as you enter here you will get here numeric.

Similarly, in case if you try to find out the mode of vector which has got 4 values 5 6 7 8 all numbers, then this comes out to be numeric, and if you try to find out the mode of our character string like as here, India this will come out to be here as a character, and in case if you try to create a vector of characters using the combined option like as here I am trying to combine 2 letters India and USA this will also come out to be as a character.

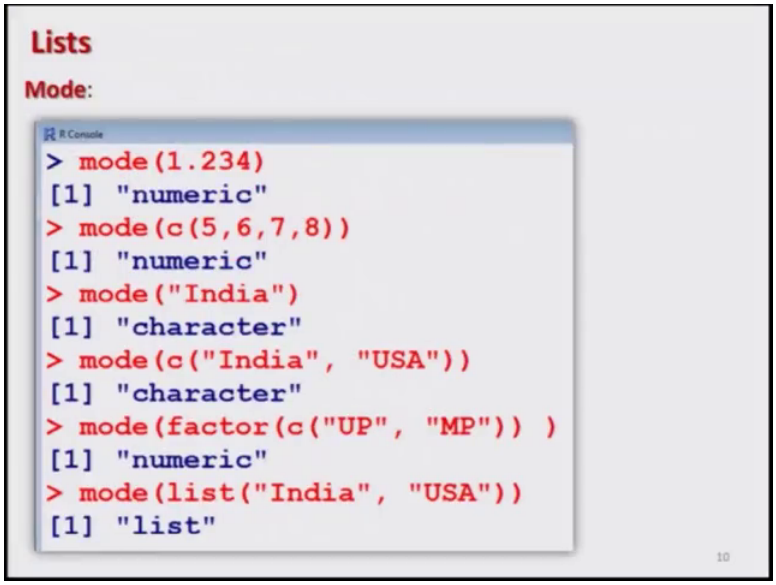
(Refer Slide Time: 20:16)



Now, first let us try to see on the r console that whether this works or not, you can see here, this is now the mode of 1.234 and this comes out to be numeric. Similarly in case if I try to find out here the mode of a vector say 3 5 7 8 and so on. Then this comes out to be a vector. And similarly if I try to find out here the mode of my name say Shalabh this comes out to be as a character.

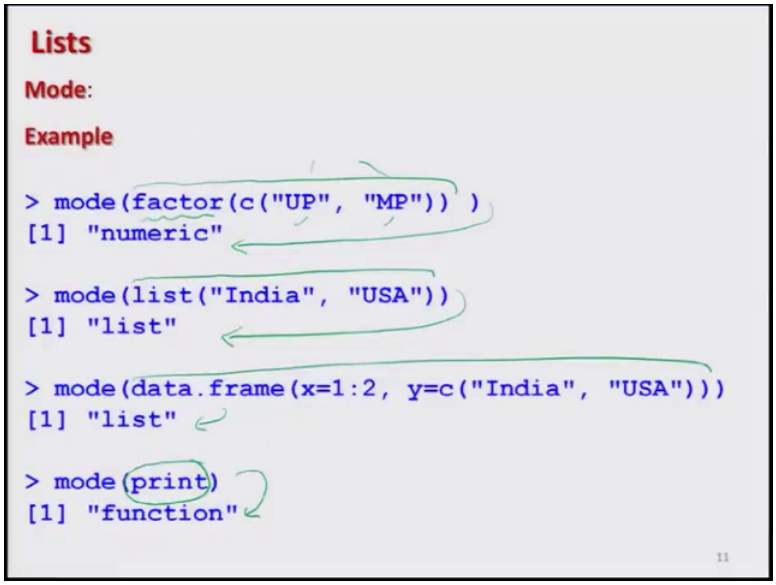
And similarly if I try to take here vector of this character say Shalabh comma this R course comma say MOOC, that we have to remember that all the, this letters have to be combined inside the c operator and we within double quotes, so you can see here this comes out to be here a character.

(Refer Slide Time: 21:26)



So, this is the screenshot of all this operation well I have quicken a different name, but the outcome is the same.

(Refer Slide Time: 21:36)



Similarly, in case if you try to take here another object like as here factor then I have to write it nearly similar to what I did in the case of say head character, but the difference is that, here I am trying to use here 2 characters UP and MP Uttar Pradesh and Madhya Pradesh, and these are 2 character strings, but when I am trying to write down here a factor, this mode is coming out to be numeric that you have to keep in mind, that is the basic difference what we did as a that when we try to take cricket here as a factor. And if you try to see here this is here also in the table also, that the difference between this character of a string, and here this thing, here I have got character, but here I have got here numeric.

So, when we are trying to find out the mode of a vector of character strings, and a factor both are containing the vectors of characters, but their modes are different that is an important point which you have to keep in mind. And similarly and I try to create here a list with 2 characters is again you will see here that this comes out to be here as a list.

And similarly if I try to take here that data frame and I try to find out its mode this comes out to be here as a list. And similarly if I try to take here a function say here print and then this outcome comes out to here as a function right.

So, let us try to do this thing in the r console, and let us try to see here what happens, say about factors you can see here this is coming out to be numeric, and similarly if you try to see here for the list this comes out to be here as a list, and similarly if you try to go for the data frame this comes out to be here as a list. And similarly if you go by this here function print then this comes out to here as a function.

So, this is how this mode function works. And in the next slide I have taken the screen shot of the same operation over here, so that you can have a look and you can practice yourself with the same command to verify whether we you are also getting the same outcome or not.

So, now at this stage I would like to stop here so that you can have some time to understand the basic fundamental off list, and to understand about different types of modes. And in the next lecture I will continue with the list, till then goodbye.