Essentials of Data Science with R Software - 2
Sampling Theory and Linear Regression Analysis
Prof. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Introduction to R Software
Lecture - 03
Calculations with R as a Calculator

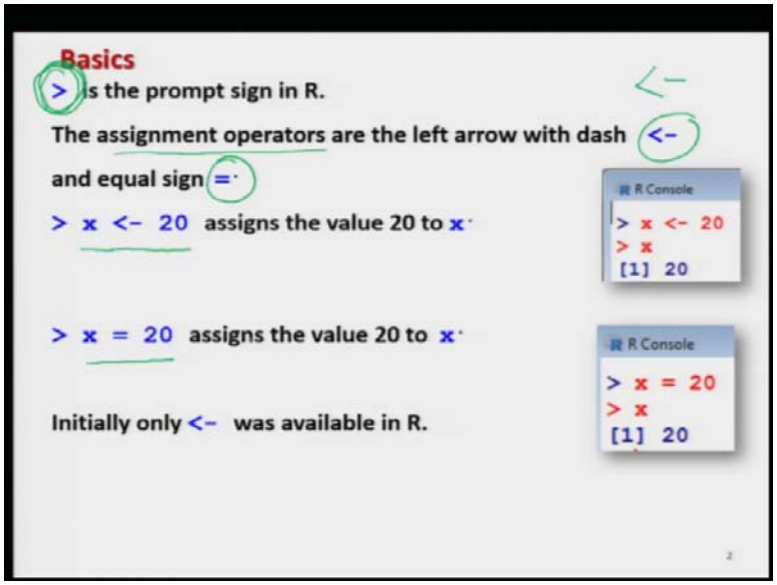
Hello, welcome to the course Essential of Data Science with R software 2 where we are going to talk about sampling theory and Regression Analysis topic and in this lecture we will continue with the topics of Introduction to R software. So, you may recall that in the last lecture I had given brief description that how are you going to start with R.

My objective was very simple; suppose if there is a student who do not know about R much, but he has a brief idea, he does not know that much details of the R software; possibly he or she will get courage to start to learn these topics, right, ok.

Now, in this lecture; I will try to take some more commands and I will try to show you that how this R can be used as a Calculator. One thing I would like to mention here that every software has its own peculiarities, the way we write any command, any function they may differ. Similarly, R also has some differences; the commands are written in a different way than in other software. I am not saying that they are extremely different, but there are differences.

So, my objective is that I want to make sure that you know what you are doing. In order to fulfil an objective, how you have to write the command that you must know that is my simple objective. So, in this lecture I am essentially going to explain how a simple calculation like addition, multiplication, subtraction, division, BODMAS etc. can be used in R. So, let us start with the course.

(Refer Slide Time: 02:33)



So, in general you know that when we start our R program; then you get here a symbol like greater than, this is here the prompt sign. And when we try to assign a value to a variable; the assignment operators, there are two operators. So, one assignment operator is you have to write the sign less than and hyphen. So, this looks like a left arrow and the dash and second option is this that instead of writing this left arrow sign, you can write simply the equality sign, right.

For example, if I take a variable x and if I want to assign the value 20 to it, I have two options that I can write down here x less than hyphen 20 or I can write down here x equal to 20; both these commands will assign the value 20 to x, right.

So, now a very obvious question crops up to our mind that why there are two commands? So, I am the one who has seen the growth of this R software. When I was pretty young means now I am no more that young. So, when I started my career in 90's and when this S plus was my first encounter and then came R.

In S plus, they used to use the command left arrow, instead of equality sign. At that time, it means I had learned different types of language like cobol, fortran, basic etc. they all were using the equality sign for assigning the as an assignment operator, but S plus was the first language which used the symbol less than and dash.

So, initially it was extremely not convenient to me at least that why they are using, but anyway that was the only software which I had to use. So, I practice it and then since S plus was a very good software, it has a capability to program and to use the built in commands, but it was an

expensive software. So, a group of people all over the world, they gathered together and they started developing the R software on similar lines I would say.

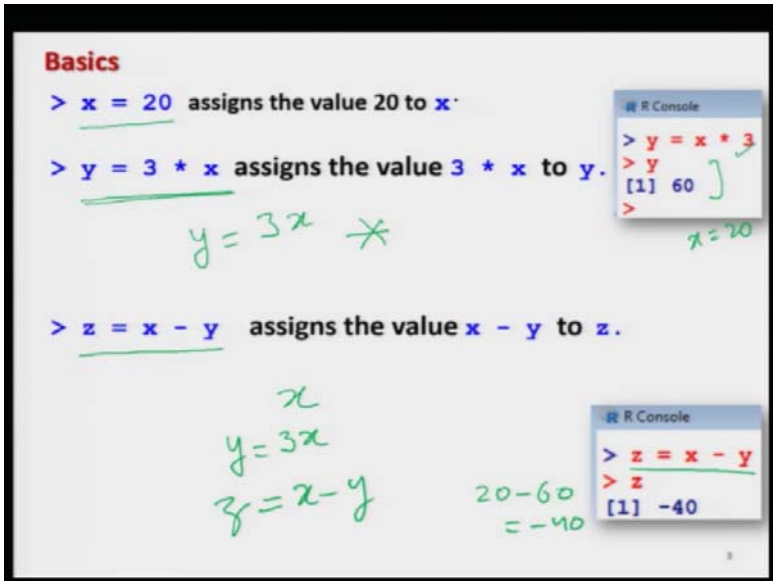
So, that is why since they were trying to replicate the S plus software; so they also use the same symbol- less than and hyphen. So, for a very long time, the assignment operator in R software was less than and dash symbol. And since I grown up with that symbol; so that has come into my habit to write this symbol as an assignment operator.

Although, I try my best to use the equality sign, but I am sorry because of my habits sometime I use this thing. So, I will try my best to use this equality sign as well as; means as far as possible, but if I use the sign less than and hyphen, you please do not get confused.

That is not your problem, means I have to change myself; that is my responsibility, but yeah, but you have to help me out so that I can take; I can take care of my responsibilities gradually with your help, I can say, right. So, means anyway; so let us try to move forward and I will try to use both the symbols and you please excuse me, if you get confused between the two, but in my opinion that will not make any difference, right, ok.

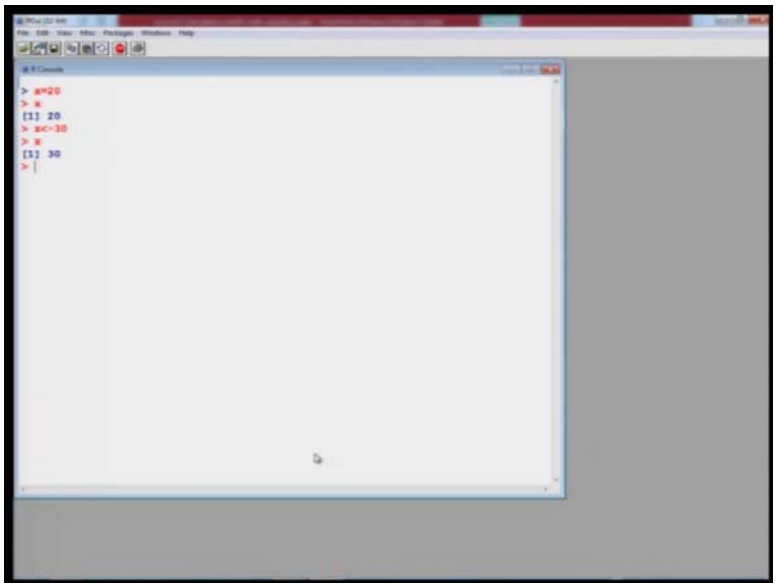
So, and then now what I will do that whatever I am showing you here, I will try to show you on the R console also so that you get more confident. Yes, whatever I have written here that is also working. So, I will be switching between my slides and the R console.

(Refer Slide Time: 07:16)



So, for example, I have shown you here that x equal to 20 is the operator and for; we need, we are using two operators equal to sign, as well as less than dash sign; so I will try to show you on the R console. So, let us try to come back on the R console.

(Refer Slide Time: 07:33)



And I can show you here for example, means if I show you here. So, you can see here I am writing here x equal to 20. And then if I try to press it here x, it will give me the value 20 and means if I try to write down here; the same simple x less than hyphen, say here 30; so you can see here, now the value of x is changed to 30.

So, now, I have shown you that this both the symbol equality sign, as well as the left arrow sign both are going to work, right, ok. Now, let me come back to our slides ok. So, now, suppose if I say here x equal let me assigned it here; x equal to 20 to a variable x, so I can use this symbol x equal to 20.

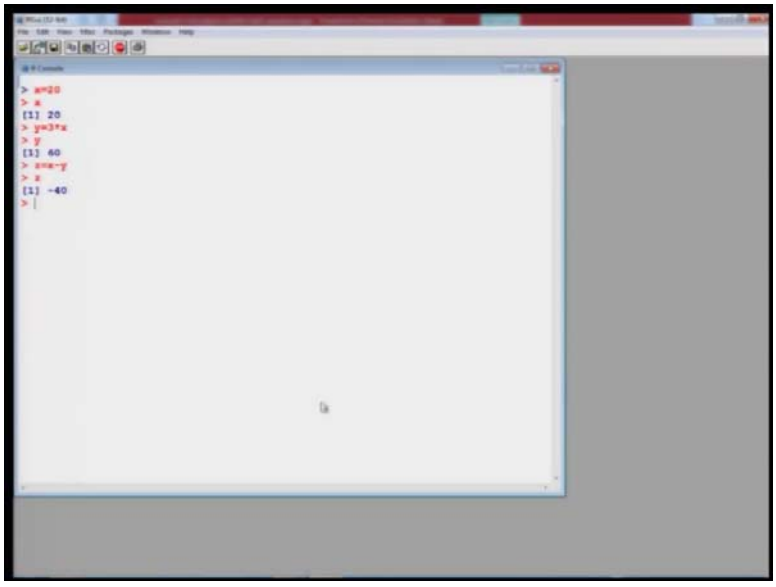
Now, just like any other language; in R also you can assign a variable to a new variable. What does this mean? Suppose, if I defined a new variable 3 into x; so now, this is my new variable and I am assigning it to here new variable here y. So, I can do it here y is equal to 3 into x; yeah. I will show you later on that the multiplication sign; multiplication symbol in R is given by this star sign which is available on your keyboard.

And suppose, I move further I have here two variables; one was here x, then I define a new variable here x; based on here this x and this new variable was here y. And suppose I want to

define here new variable which is based on x and y; suppose I define here it x minus y and I call it here as a z.

So, I try to define here a new variable here z which is equal to here x minus y and I try to execute it on the R console. I can show you here that I am using here x star 3, this will give me the value here x equal to 60 because x is here 20. So, y will become 3 into 20 that is 60 and if I try to make it here z is equal to x minus y; so this 20 minus 60 that will become here minus 40. So, I will try to show you it on the R console also.

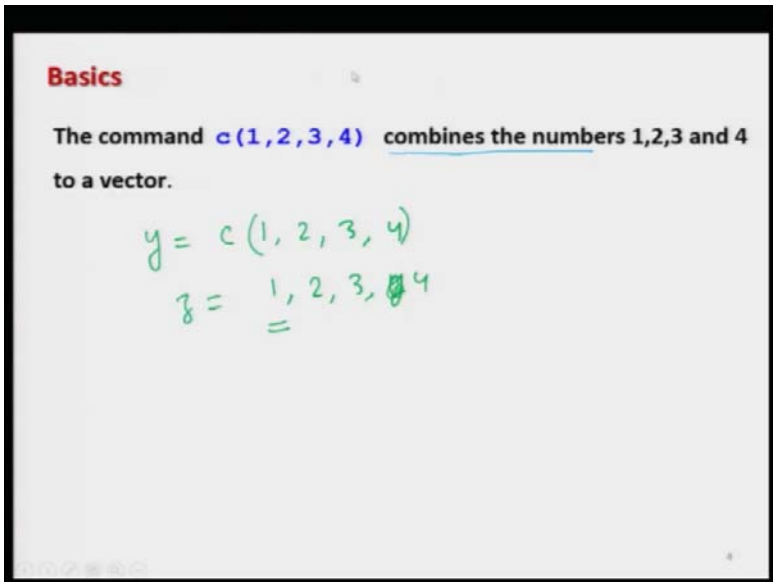
(Refer Slide Time: 10:14)



So, I can show you here; suppose I want to clear my screen, so as I said I will press here control and l. So, control and l and x equal to here 30; so, I try to take it here x equal to 30; well, I have taken here x equal to 20. So, let me take it here x equal 20; now x equal to here 20, y is equal to 3 into x; that is here y.

So, you can see here 3 into 20 here is 60. Now, I try to define here another variable here; z is equal to x minus y and you can see here the z value here is coming out to be -40. So, now you can see here; this is what I was trying to convey that in R just like any other software; you can store a scalar or a variable in a variable, right, ok.

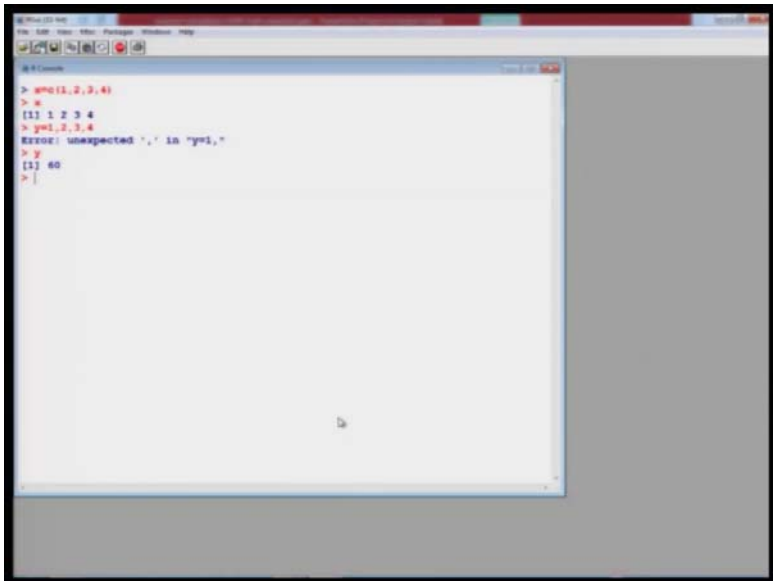
(Refer Slide Time: 11:03)



And then we have a command here which is denoted by here c; this commands here, c will combine the numbers that is used for combining different types of numbers. For example, if you have a set of numbers say here 1, 2, 3, 4 and if you want that R should read all the numbers, then you have to write it here c and then inside the parenthesis you have to write these numbers separated by this comma.

Well, in case if you try to write down only here 1, 2, 3, 4; then you will see that there will be an error and in most of the cases the R will read only the first value. For example, I can denote it by here y equal to c 1, 2, 3, 4 and if I try to define here z equal to 1, 2, 3, 4; this will have different types of this numbers. So, let me try to show you what do I mean by this thing.

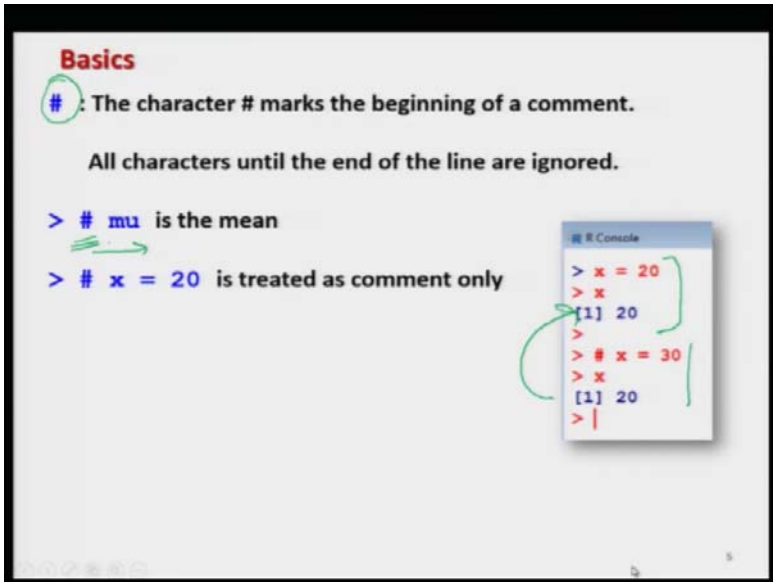
(Refer Slide Time: 12:13)



So, if I come here R console, you can see here; now I clear my window and I try to write down here x is equal to suppose here c 1, 2, 3, 4; they have to be separated by commas. So, now you can see here now it will read here c equal to 1, 2, 3, 4. But now suppose if I say here; if I give here y is equal to 1, 2, 3, 4; now you can see here this is giving me an error and if I try to give here the value of here y, it is not giving me the value of here 1, 2, 3, 4, but it is giving me the value of here y which I had used earlier.

And if you try to recall in one of the earlier lectures, I had requested you, in the last lecture actually that whenever you start a new programming language; you need to remove all these variables, otherwise that will create a big confusion, right.

(Refer Slide Time: 13:08)

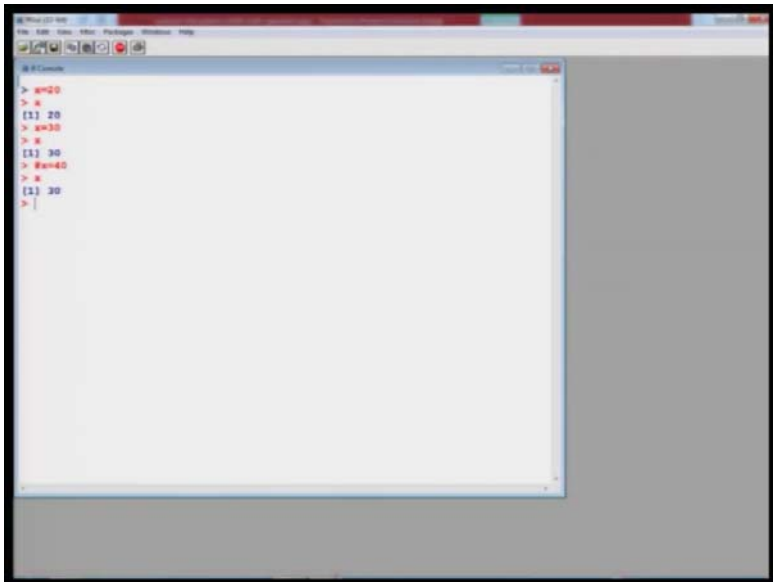


So, anyway, right; similarly, while doing the programming; if you want to write a command as a comment; that means, you do not want that the command to be executed, then we use the symbol here hash. Yeah, now a days hash is a very popular symbol, hashtag and these are new terminologies which have come.

So, means I am sure that all the young generation knows about hash, right; so, this is the symbol here hash. So, if you try to write down any command in which the first letter is hash, then anything whatever is written after hash will not be executed and this will be treated as a comment only.

For example, if I try to do it; I will try to show you on the R console, if I try to write down here x equal to 20; then it will give me x equal to 20. But, if I write down here hash; x equal to 30, it will again give me the same value; it will not consider x to be 30 right. So, I will try to show you it on the R console.

(Refer Slide Time: 14:29)

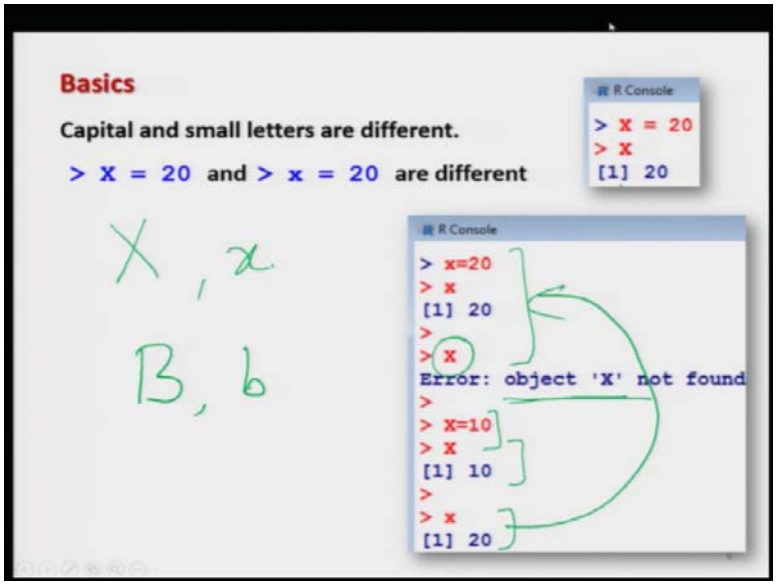


So, now, if you try to show you here that this is here x is equal to 20. So, I can show you here that x is coming out to be 20 and now if suppose if I take here x equal to 30, then now if you see here values of x; this will come out to be 30, but means if I write down here hash x is equal to 40, you can see here still the value of x is same 30.

Once when you take it; took x to be 20, then you took x to be 30, then 20 was replaced by 30. But when you took this hash x equal to 40, then means again this was not a replacement; the

value of x equal to 40 was not really assigned to x, but the earlier value 30 was considered right.

(Refer Slide Time: 15:24)

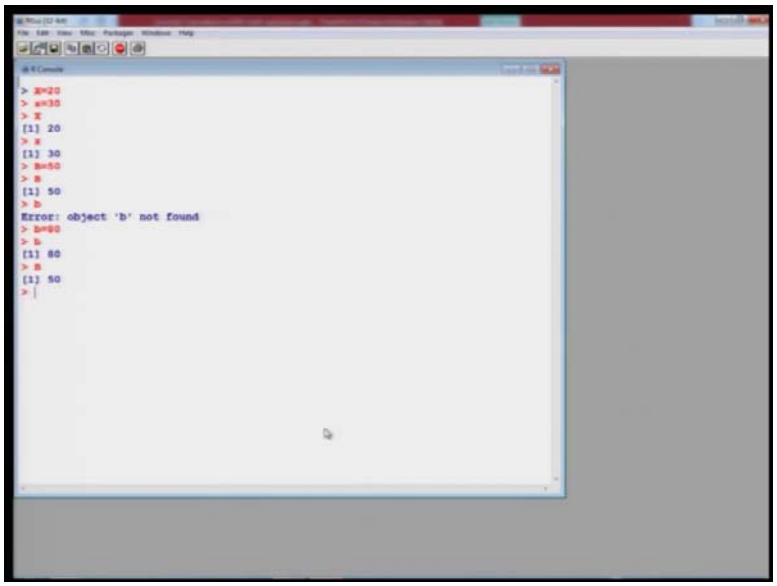


So, now the next very important thing that in R; this capital and small letter; they makes a huge difference, they are just taken as two different values. For example, if you write down here capital X and say small x, these are two different things, right.

Yeah for x it might not be so clearly visible, but if I try to take it here see here a variable name here is capital B and say small b; they will be clearly visible and I will show you that these are not the same thing. For example, if I try to take here capital X equal to 20, then and for the small x equal to 20 and if I try to show you here the value of here X; capital X, then it will show that there is an error; it is not there.

But, if I try to take here capital X to be 10, then it will give me capital X to be 10 and small x to be 20 which is coming from here. I will try to show you it on the R console which will make you more clear about these things, right.

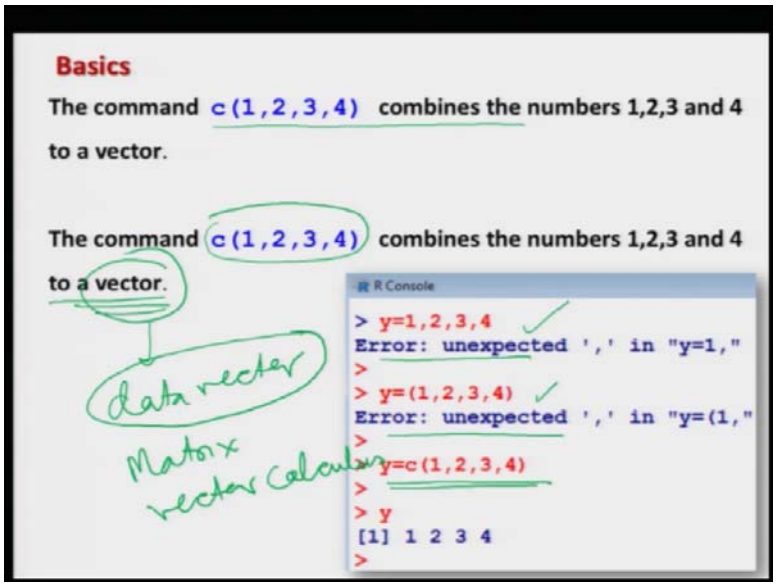
(Refer Slide Time: 16:28)



So, if you try to see here; now if you try to see here, if I try to take here x equal to 20 and say small x equal to 30; you may not see the difference on the screen. So, but anyway since I have taken; so I will try to show you capital X is 20 and small x is 30, but I will try to take it a variable here capital B. So, B is going to be suppose here 50; so you can see here capital B is 50, but if I try to take here a small b equal to here 20; no first, I try to show you what is it the value of a small b; you can see here b is not found.

Although, you have given the value of capital B; so now, let me take it b is equal to here say here 80; small b equal to 80. So, now you can see here small b is 80 whereas, capital B is again 50. So, a small b and capital B; these are the two different variables which are assigned to these things, right, ok.

(Refer Slide Time: 17:26)



So, yeah, this I already have shown you; so I will just give you a brief thing that this command c 1, 2, 3, 4 is used to combine the numbers. And one thing which I, the reason why I am trying to repeat it here, I want to make you careful here that this command c 1, 2, 3, 4 combines the number into a vector.

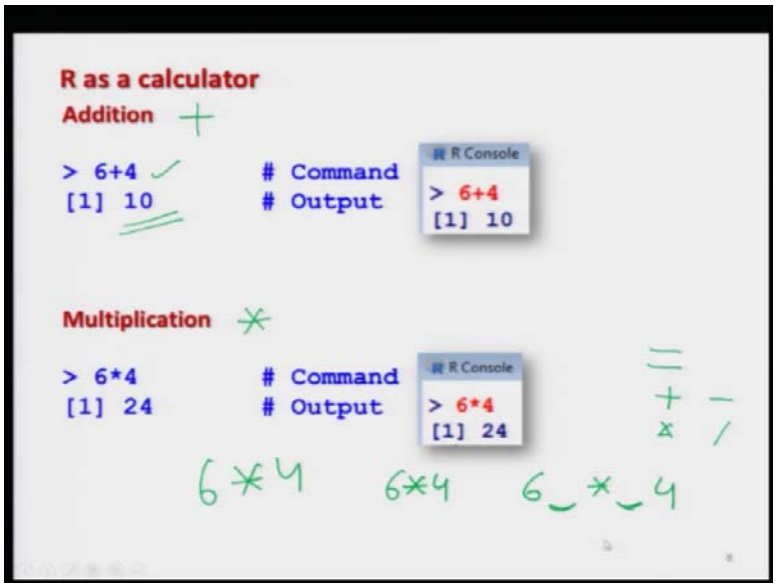
Now this becomes sometimes very confusing and that is why I have written it two times here the same sentence so that you means; I will be putting my pen of different colors on the second line; so first line will always be clear to you.

Once you say vector, vector is a very confusing thing, there is a data vector, data vector means simple vector, just try to combine all the things. There is a vector from the matrix theory and beside those thing; there is a things which are there in vector calculus also. So, these are different things, but they are all called as vector. What I would like to point out here; that when we are trying to combine the data here in c 1, 2, 3, 4, we will also call it as here vector but this is essentially a data vector.

The operations of matrix theory will not be operated on such a vector and if you want to have a vector like a vector in the matrix theory; then there is a different command matrix, then we will that we will try to discuss later on. So, that is my idea here.

And I already have shown you this outcome that if you try to go give here 1, 2, 3, 4 or inside the; parentheses also 1, 2, 3, 4, it will give you error; it will not work unless and until you write y equal to 1, 2, 3, 4 with the c command; so, that is my idea of this slide.

(Refer Slide Time: 19:27)



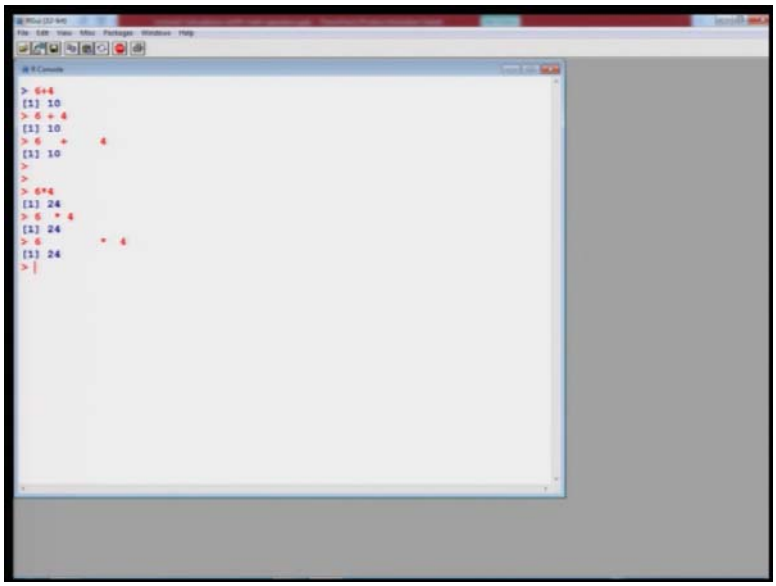
Now, I come on another aspect that how you can use this R as a calculator, how you can operate the different types of commands like as addition, subtraction, division multiplication.

So, in R; if you want to use the addition sign, then the symbol is very simple, the usual plus sign. So, if you try to give it here 6 plus 4, then this is here 10; means if I want to add 6 and 4, I have to simply write 6 plus 4. And if I and similarly for multiplication, the symbol here is star. So, if I want to multiply 6 and 4; I have to write down say 6 star 4.

And yeah I just want to make it clear that there is no issue of the blank sign between the two number. For example, if I want to write down here 6 into 4, I can write down here 6 into 4; just continuous without any blank space between these three values or I can write down here 6 blank space star, some blank space and here 4; they will be, they will be giving us the same number.

And this I would like to clarify that this is true for all the mathematical operators; either this is equal to sign, addition, multiplication, later on we will do subtraction, division; whatever you want this blank sign is not really going to make a difference.

(Refer Slide Time: 21:06)

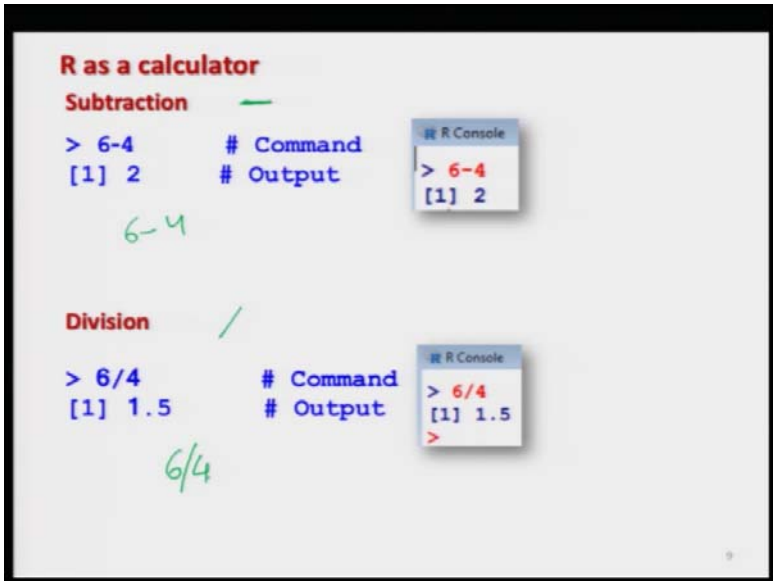


So, let me, and I try to first show you these things on the R console and then I will move forward. So, if you want to add here two numbers; suppose if I want to write here 6 plus 4; 6 and 4, I want to add; you can see here this is 10.

Now, you can see here in this 6 plus 4; I have not given the space between the two. So, I now I try to write down here once again 6 with the space, here 4, this is here 10. And now if I try to add here more space; even then you will see here the answer is 10. So, blank space has no value, this is what I want to convey.

And similarly, if you want to multiply here two thing; 6 into 4; so you have to write 6 star 4 which is here 24. And even if you want to give here blank space between the two; unequal blank space whatever you want; so, you can see here 24, even if I try to make it here something like this 6 and 4 means unequal number of blank space even then it is 24. So, be assured this is; this blank space not going to contribute anything to us, yeah; unless, until I tell you right.

(Refer Slide Time: 22:10)



Similarly, if you want to subtract anything; suppose I want to subtract 6 and 4, then I have to use the symbol here; the classical traditional symbol minus right. So, means I will; so if I want to add here 6 minus 4, so this will give me an answer 2 and if I want to use division; then I, then the symbol here is this backslash, right.

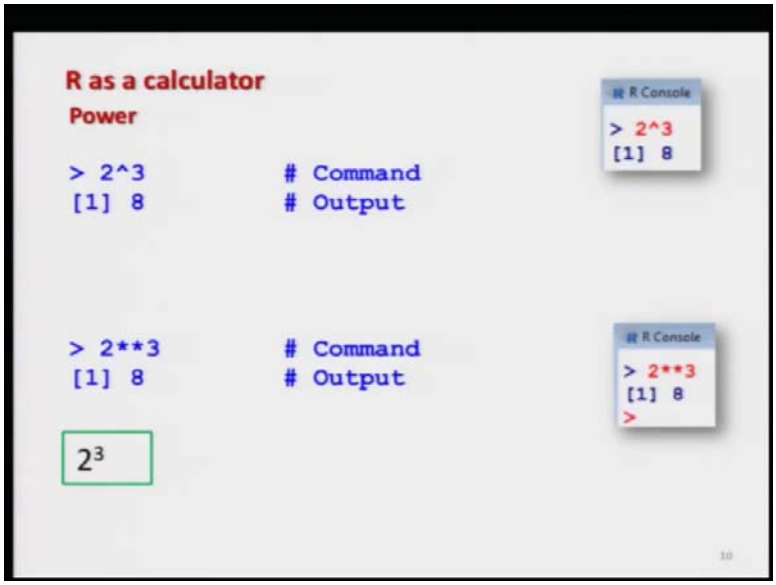
So, for example, if I want to divide 6 by 4; so I have to write 6 backslash 4, it will give me the answer 1.4. One thing, I would like to point out which is a very common mistake because I will be providing you the slides and in many cases and also in assignments also; what will happen; that; in many cases, you will try to copy and paste the commands from my slides or from my notes or from the assignments.

These syntax, these slides and these assignments have been prepared in software like MS Word, Power Point, etc. This MS Word has a very peculiar property that once you write 6 minus 4, then as soon as you type 6, then minus and then blank space and then 4; you will observe this minus sign is converted into another symbol which becomes larger.

You can experiment it; that if you write 6, blank sign, blank space minus blank space 4. As soon as you type 4, this hyphen which is the minus sign becomes bigger and it becomes something else. So, by looking at the eyes; by looking with your eyes, you may not be able to identify whether this is minus sign or something else. But when you try to copy the entire command and paste it on the R console; that may create some trouble and it will give you there is an error in the command or some error message.

In those situations, you have to be very very careful that when you are trying to use the minus sign; try to type the minus sign inside the R console itself. This will not happen in case if you are trying to copy and paste the command from the R studio, R script window; that will not happen. So, this is a point where sometime I see this that student get confused and they waste their time; so, just be careful, right, ok.

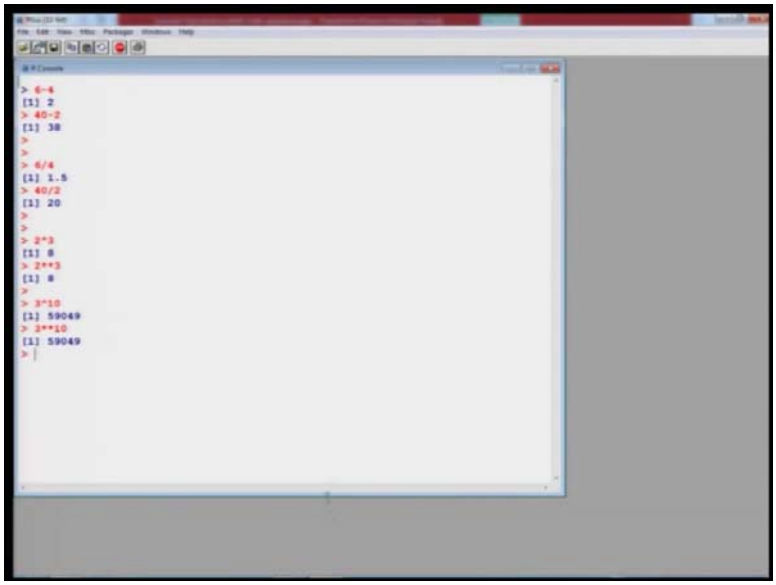
(Refer Slide Time: 25:05)



So, now I come to my slides back; now in case if you want to use the power operator, power operator means suppose I want to write down here 2 raised to the power of your cube. How to write this thing?

So, for that I have here two options; first option is I can use here the command hat and second option is I can use the command here 2 star; double star. So, I can write down 2 cube here like I said 2 hat 3 or 2 star star 3 and they will give us the output here; see 8, right. You can see here, this is here the screenshot; let me try to show you it here on the R console.

(Refer Slide Time: 25:55)

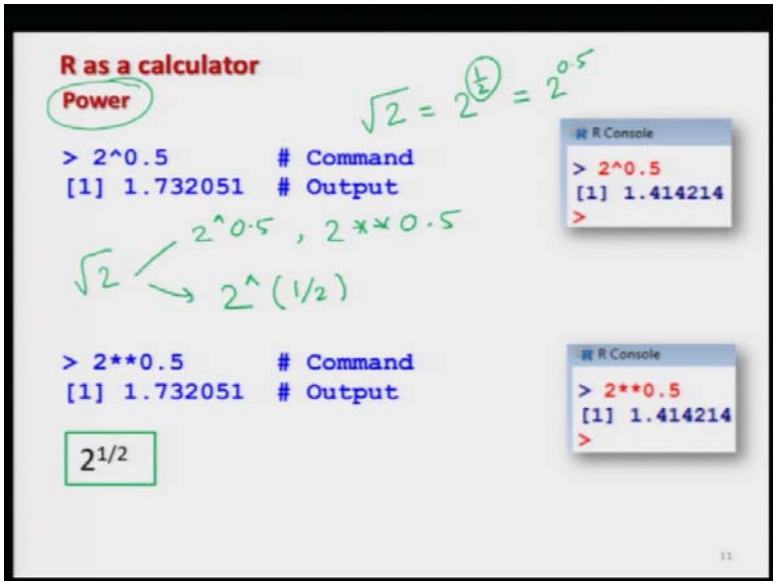


So, so first let me show you here the difference part. Suppose, if I want to subtract here 6 minus 4; so this is 6 minus 4 and here 2. And similarly, if you want to here 40 minus 2; this will be here 28 and suppose if you want to divide this will be here 6 divided by suppose; if you want to divide 6 by 4; so this will be 6, backslash 4 like this and answer will be 1.5.

And similarly, if you want to divide here 40 divided by 2, the answer will be here 20 and similarly if you want to write down here 2 cube. So, you will have here two options either you write 2 hat 3, this will be here 8 or you write 2 star star cube which will be over here 8.

Similarly, if you want to write down here 3 raised the power of your 10; this is 59049 and if you want to write 3; raised to the power of 10 at 3 star star 10, this will be something like the same answer 59049. So, you can see here that these are the different ways by which you can write down these things, right.

(Refer Slide Time: 27:00)



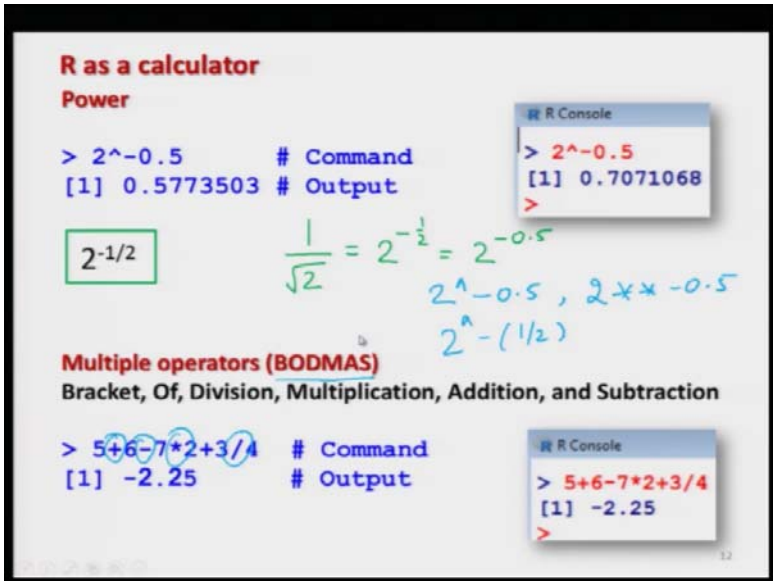
And yeah, similarly if you want to write down here the square root; if you want to find out the square root, so you know that square root of say here 2, this can be written as here 2; raised to the power of here 1 by 2.

So, now I simply have to learn how to write down here this fraction 1 by 2. So, this is; so there are two options, either you write down here as a 1 by 2 which I will discuss later on that how to express this thing or you can write down here 2 raised to the power of here; 0.5, right.

So, I have here two option; if I want to say obtain here square root of 2, first option is 2 raised to the power of here 0.5. Or second option is 2; star star, 0.5, our third option is this that I can write down here 2 raised to the power of here, inside the bracket you know that much mathematics I can write down here 1 by 2, right.

So, I will try to show you that all these things are going to work, but the moral of the story is that, that if you want to use the power of a number, you want to find out the power of a number. So, that is valid for integer as well as for fraction and you have to use it in the exactly same way, as you have done it earlier.

(Refer Slide Time: 28:27)



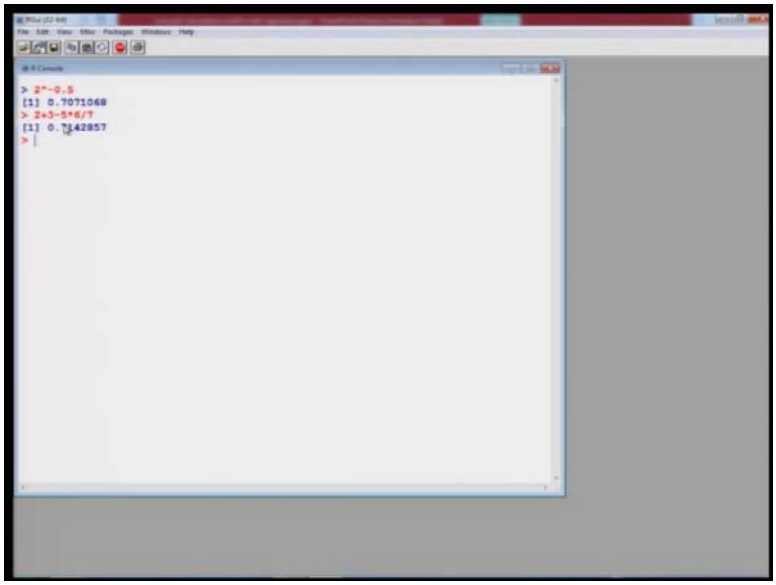
And similarly, if you want to use here; you want to find out here 1 upon square root of 2; 1 upon a square root of 2 is same as 2 raised power of here minus 1 by 2 which is same as here 2; 2 raised power of here minus 0.5, right.

So, you can write down here; here as say 2 raised to the power of here minus 0.5 or you can write down here 2 power of here; sorry, 2 double star say here minus 0.5 or 2 hat minus 1 by 2 and so on; whatever you want, you want, you can use it.

And in case, if you have this multiple operators like as you want to use say this here plus sign, minus signs, multiplication sign, division sign; everything in one command, then the rule is very simple, we are; the R also follows the rule of BODMAS; that is the Bracket, Of, Division, Multiplication, Addition and Subtraction; that is the same what you have learned right from your childhood right.

So, let me try to show you all these things over here; it means I will try to show you here this means power operators and this BODMAS rule, ok. So, let me come to here this thing.

(Refer Slide Time: 29:53)



So, power operator; I already have shown you here. So, suppose I want to find out 1 upon square root of 2. So, this I can write down here 2 hat, say here minus 0.5; so you can see here, this is the same thing, right. And if you want to write down here 2 plus 3 minus 5 into 7 into 6 divided by 7 and so on; so in these cases the BODMAS rule will be followed and you will get an answer, right, ok.

So, now in this lecture; I have shown you how this R can be used as a simple calculator and you can obtain such results, right. These are not very difficult things and these are the thing which are usually the same as what happens in other software, another languages, but you need to practice it.

You must know that if you try to write down a complicated expression where you have addition, multiplication, subtraction; what is really happening, that is very important for you. Because whenever you are trying to do the programming; one of the very important step is that before we execute the program on a real data, we always try to take a very small data set and we try to make all the calculations manually with our own hands.

And we want to make sure that the program is doing the same thing; what we are asking the program to do it. So, in order to understand those things, you need to understand these smaller things very clearly. So, I will say try to take some simple expression yourself, try to write down on a piece of paper, try to solve them with your mathematical knowledge and try to type them on the R console and verify whether you are getting the same outcome or not.

This will boost your confidence so that when you are writing a big program, then you will be more confident. So, you practice and I will see you in the next lecture with some more commands on R.

Till then good bye.