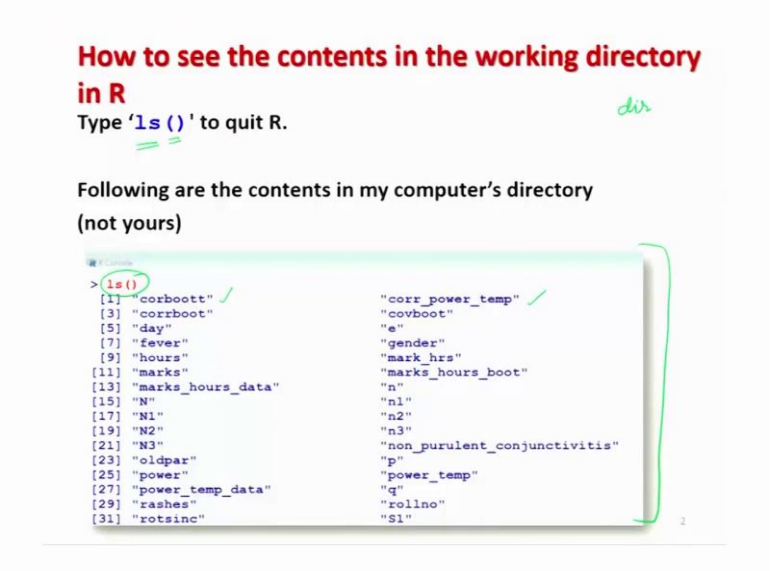
Foundations of R Software
Prof. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Introduction
Lecture - 06
Basic Operations in R

Hello friends, welcome to the course Foundations of R Software and in this lecture we are going to talk about some very basic, fundamental, elementary operations in the R software. You know this course is a very basic course. So, there are many operations in the R software which are needed, which are very basic fundamental, which are not difficult at all, but they are needed to understand the programming and structure of programs and when we are trying to do programming those actions are needed.

For example, in case if you want to clear the screen or if you want to see the directory or if you want to set the working directory, etc. These are very simple things, but these are the basic ingredients of any programming. So, in this lecture and in the next lecture we are going to understand these very minor basic elementary operations. So, we begin this lecture and I will try to show you first on the screen and then I will try to show you on the R console. So, let us begin our lecture ok.

(Refer Slide Time: 01:31)

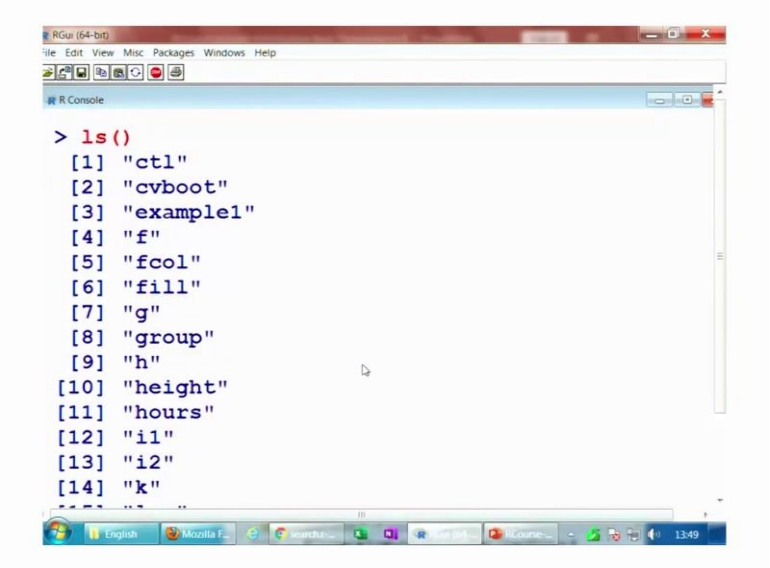


So, one of the basic operations when you are trying to work in the R software is that, you would try to see what are the files etc. which are available in your working directory. For example, in case if you are working in MS-DOS then there is a command like dir, which tells you about the contents of this directory.

Similarly, in the R software, we have a command here ls and if you write l s and this parenthesis, then you will get the contents of the working directory in your computer not my computer remember. For example, you can see here I have used here this command ls and I have got this outcome.

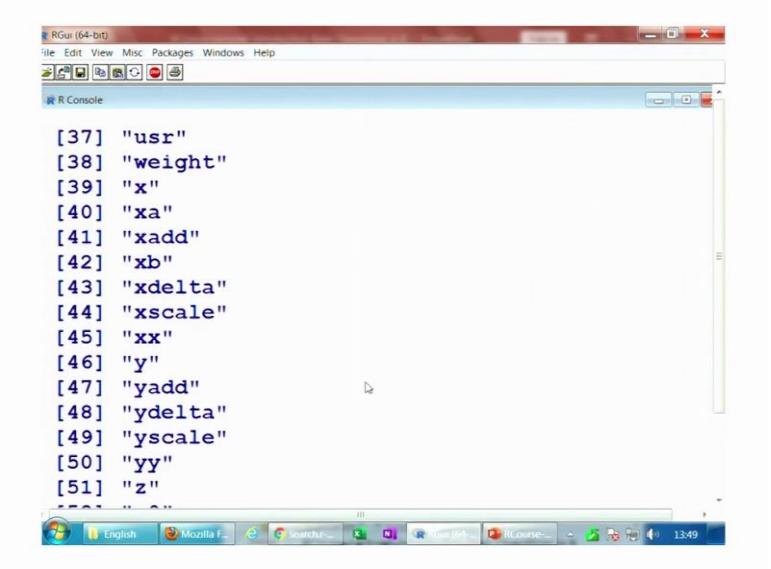
So, this is giving me that ok I have something like “corboott”, “corr power temp” etc. Well these are the program which I am using. So, they are available in my computer where I prepared these slides. Now, in case if you try to do the same command in your computer possibly you will not get the same contents, but you will get the contents and that is our basic objective.

(Refer Slide Time: 03:04)



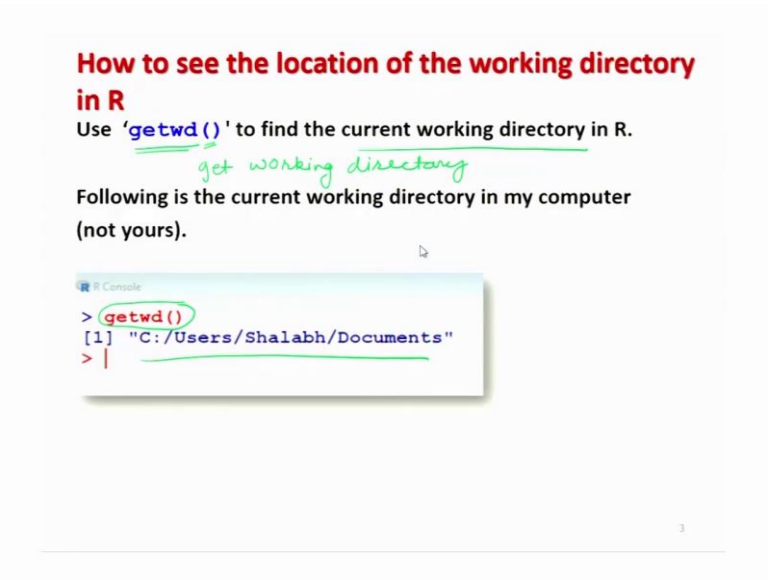
So, in case if I try to show you here this thing on the R console itself you can see here if I try to press here ls and this parenthesis you will see here these contents that.

(Refer Slide Time: 03:13)



We have here with this type of data. So, this is the files which is available here, ok. So, now, we come back to our slides and we try to understand some more basic operation.

(Refer Slide Time: 03:35)

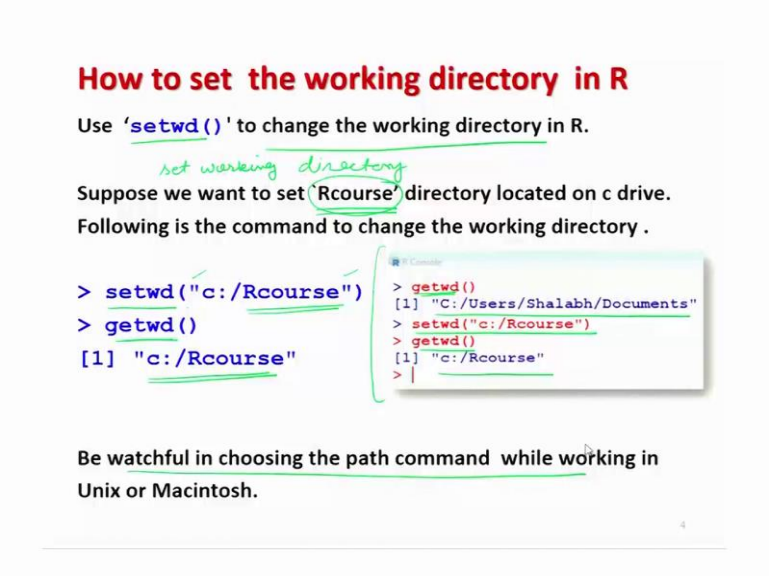


Now, suppose you start working in the R software and you would like to know that what is the location of my working directory is a place where all your files etc. they are located. So, in order to know the location of my working directory, we have a command here getwd.

This means get working directory. And then you have to write getwd and then parenthesis. So, in case if you try to execute this command on the R console, for example, if I write here getwd then I get this outcome. Well, this outcome is going to be different when you are trying to execute it on your computer, because this is the location in my computer in my laptop.

But you can see here the location is "C colon backslash Users backslash Shalabh backslash Documents" etc. and in case if you want to change it.

(Refer Slide Time: 04:56)



Suppose I have created a folder somewhere where I have copied all my files and I would like that all my files as well as the output they should be stored in that particular location in that particular directory. This means, we want to change the working directory. The question is how to get it done. So, for that we have a command here setwd. This means set working directory and after this all your output will be stored in that particular directory.

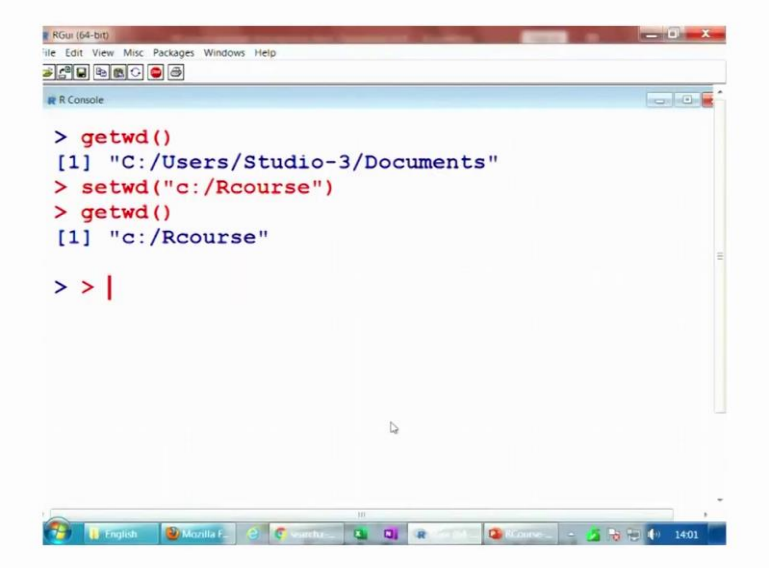
Your input data will be read from this directory and whatever you are going to do in the R software that will be executed from this directory. So, just for the sake of understanding. I have created a folder Rcourse capital R and then c o u r s e this is the. So, I have created this folder on the c drive and I want that my R software should read the files from this directory and all my output should also be stored in this directory.

So, in simple words, I want to change the working directory to be the folder Rcourse on the c drive. So, for that I write down here command setwd within parentheses within double quotes. I will write down the c colon backslash Rcourse and after that if I try to see the working directory for that I use the command here getwd I get here this new address.

One point I would like to make here that ok, I am working at present in the windows system, but in case if you are working in any other system, you have to follow the rules to write down the path in its own way, the way it is required in that platform. So, you can see here the output on the R console goes like this.

If I try to do here getwd it is giving me this directory that is c colon backslash Users backslash Shalabh backslash Documents and after that I change the working directory and after that if I try to get the working directory it is showing me c colon backslash Rcourse, right. So, you have to be watchful when you are trying to choose the path command while working in Unix or Macintosh but before going further let me try to execute these commands on the R console also.

(Refer Slide Time: 08:07)

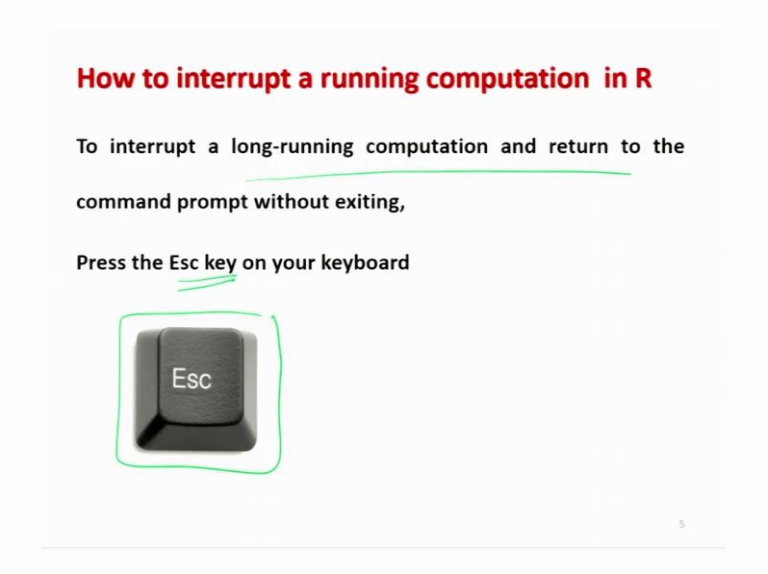


So, you can see here that if I try to write down the command here getwd this is getting the working directory. So, you can see here this is the c colon backslash Users Studio-3 Documents. Well not the same computer on which I have appeared my slide. So, this is

the computer where I am recording it. So, this is the address. Now I try to write down here setwd, I would like to change my working directory.

So, you can see here. In case if you are not getting any message or error; that means, the job is done and now in case if you try to see what is the working directory you use the command getwd and you can see here now this is c colon Rcourse, right, ok.

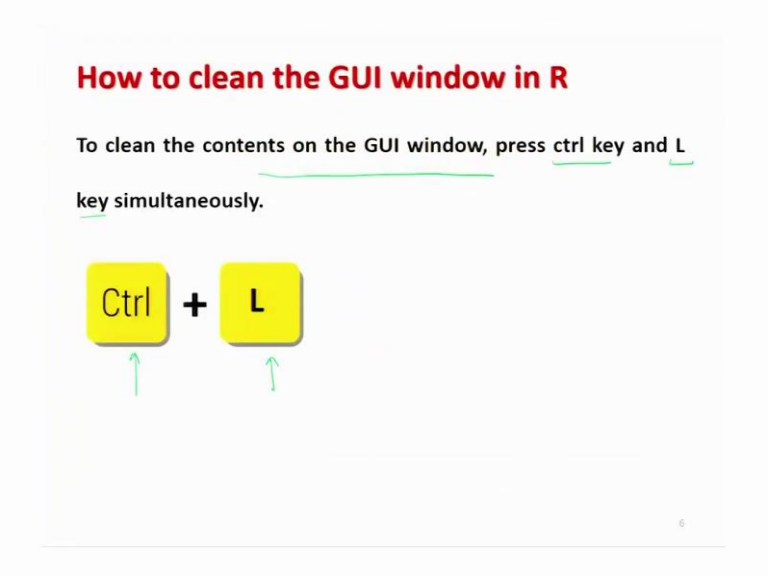
(Refer Slide Time: 08:59)



So, now, we come back and we try to understand some more basic operation. Well, this is another very basic fundamental requirement when you are trying to execute a program. Sometime you execute a program and it is working and you want to interrupt it, you want to stop it. Suppose you start a computation and it takes some time. Suppose the computation is going on somehow you realize that ok, there is something, some reason, some mistake because of which you want to stop it.

For that you simply have to press the escape key. So, you know this escape key is present on your keyboard right. So, this is very simple command.

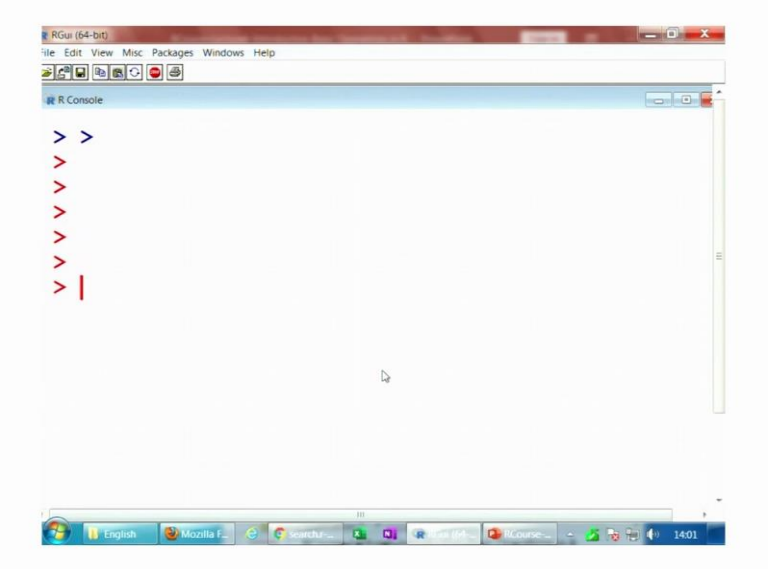
(Refer Slide Time: 09:38)



And yeah I already have talked about the next command couple of times and you know, but still just in order to complete it I will explain you here again that the question is how to clean the GUI window in R. So, in order to clear the contents on the GUI window, Graphical User Interface window, we simply have to press two keys together.

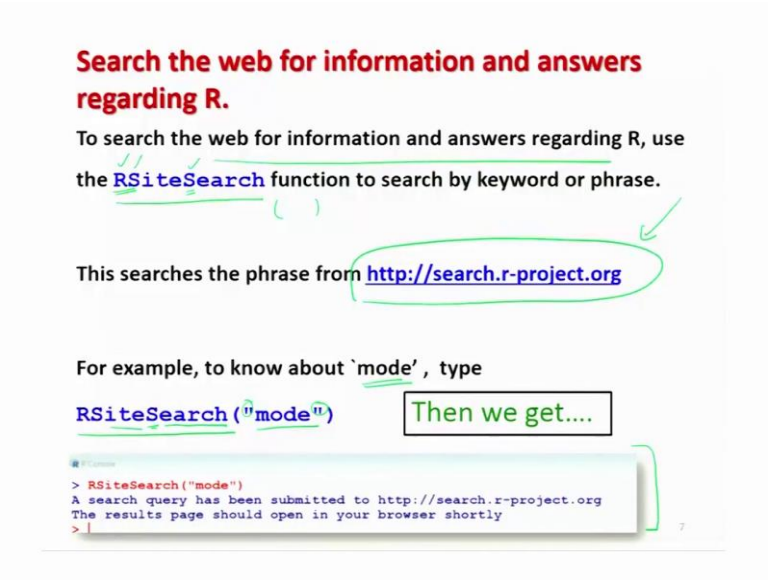
One is control key and the key with L alphabets on the keyboard. So, if you try to put here control plus L your screen will be clear right. For example, if you want to show it, although we have done it many times, but still I would like to show you. Suppose I want to clear this screen, this is the RGUI window. So, I simply press here control that is the key where it is written c t r l and then I press here L.

(Refer Slide Time: 10:39)



You can see here as soon as I do it will become clear. Suppose if I have entered and I am here and I want to go on the first line, I will again say control L and I will come back to the first line. So, these are very basic operations but they are needed.

(Refer Slide Time: 10:55)



The next command is that suppose I want to search some information on the web. So, the one option is that you can go to any search engine and type the question over there and get it done, but within R also we have a provision. And in order to search the web for

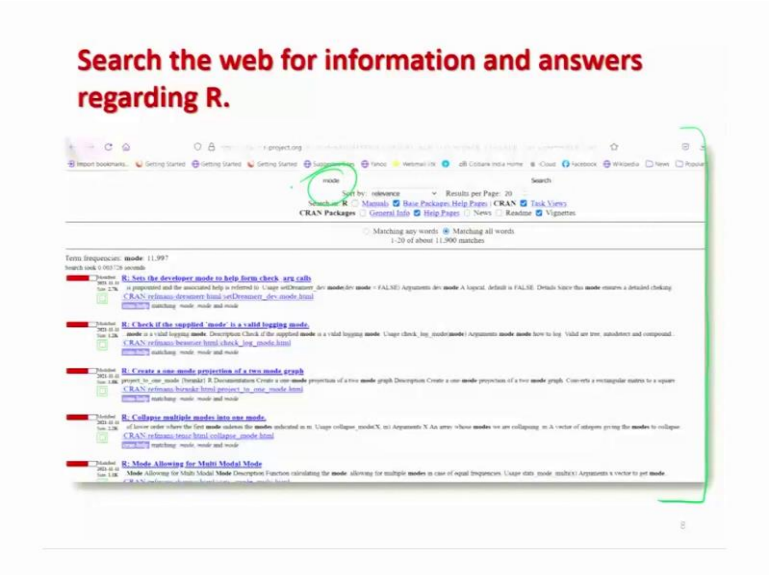
information and getting the answers regarding R, we can use the command here RSiteSearch.

Now, you have to be very careful when you try to read this command. Here this R and this S, they are in capital letter and this S of this search this is also in capital letter and all other letters are in small case, lower case. So, once you try to write down this function then inside the parenthesis you have to write what exactly do you want to find.

And then if you try to type this command on the R console, the control will go to this site; search dot r hyphen project dot org. And it will try to search for that word on this website. The difference is that if you try to go to any search engine like for example, Google and if you try to look for that word possibly it will give you many many websites where those things are available, but when you are trying to do it from inside the R software it will be restricted only to this sight.

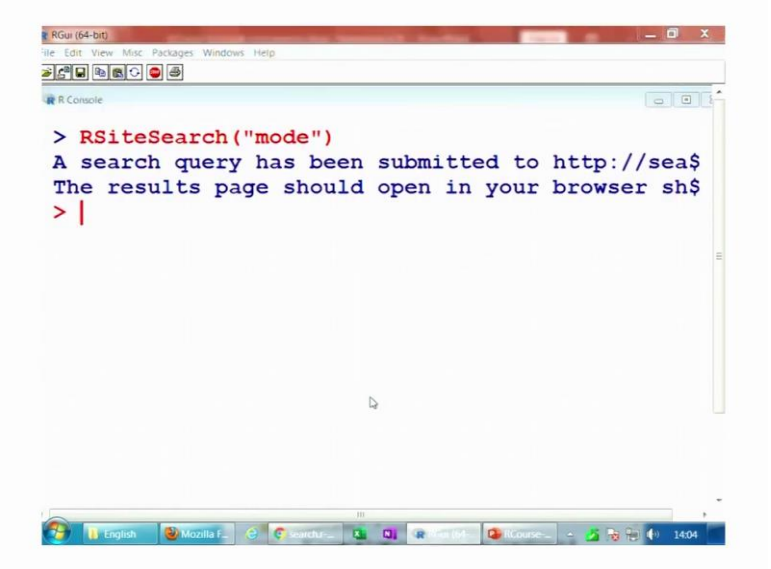
For example, in case if I want to know about mode m o d e, I will simply type on the R console capital R capital S then small letters i t e and then capital S and then e a r c h in lower case. So, in case if I try to write RSiteSearch and within parentheses within double quotes I write mode m o d e then you will see here that we get a screen here like this one and after that browser is automatically opened.

(Refer Slide Time: 13:22)



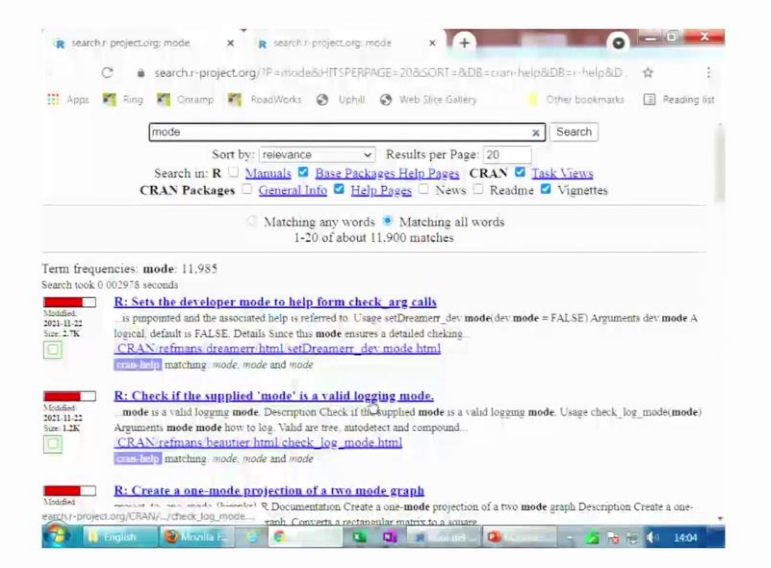
So, let me try to show you here this thing and you will see that this type of web site is open from there it will try to give you the different information which is related to the word more “mode”. You can see here this is here mode. And it is trying to do it. So, let me try to show you this thing on the R console also. So, I try to copy this command so that I can save some time and I do not make any mistake.

(Refer Slide Time: 13:56)



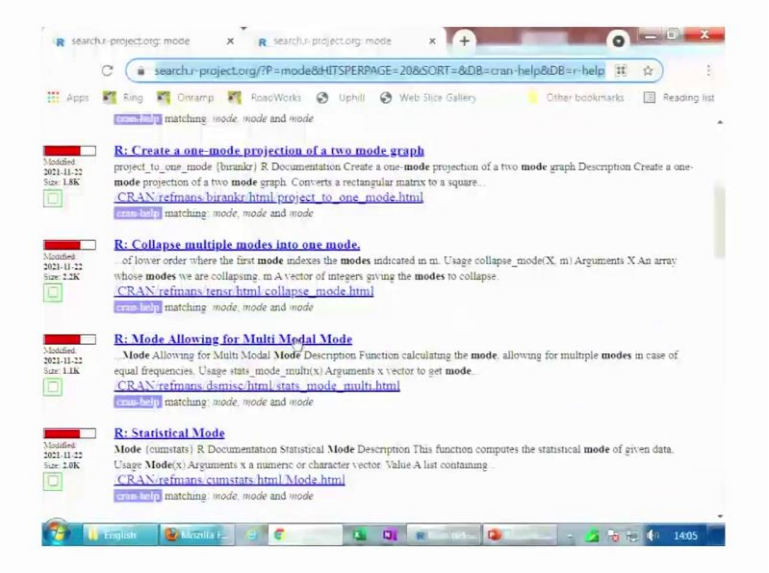
So, you can see here I try to paste this command.

(Refer Slide Time: 13:57)



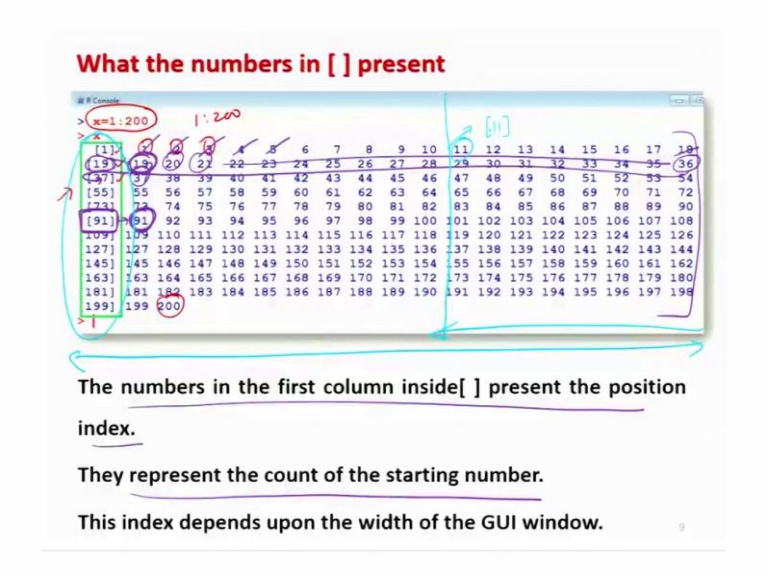
And then as soon as I do it, it will come to us website like one, you can see here.

(Refer Slide Time: 14:08)



So, now, there are various information you can say multimodal, one modes one mode projection and then now you can just look at it and in the R console it will be like this what I shown you in the R console right and the screenshot also here, right ,ok.

(Refer Slide Time: 14:32)



So, now after this I come to a very interesting issue which many times many people ask me and yeah these are very common things whatever I am telling you, these things looks

very simple very common, but when somebody is trying to learn this R software for the first time, these are the questions which crop up to their mind and they would like to be clarified otherwise that will always be a confusion in their mind.

Suppose I take an example here and then I explain you what I want to explain you. Suppose I give here a command x is equal to 1 to 200 one like this, 1 colon 200. So, you will see later on in the further lectures that through this command, I can get here the numbers from 1, 2, 3, 4 up to 200. So, you can see here this is 1, 2, 3 and then up to here 200.

Now, if you try to see in the first column I have here a green box also. There are numbers, like here this is 1 this is 19 then 37 then 55 etc. What these numbers are really indicating? Actually the answer of this question is very simple. These numbers in the first column they are indicating the index or the position of the first number in that line. They simply represent the count of the starting number.

For example, if you try to see here this is 1, 2, 3, 4, 5 up to here 18 and now the next number here is 19. So, this 19 is indicating indicated here. Then it will continue 20, 21 etc. up to a 36 and then if you try to come here it will become here 37 and this is the 37 is the first number.

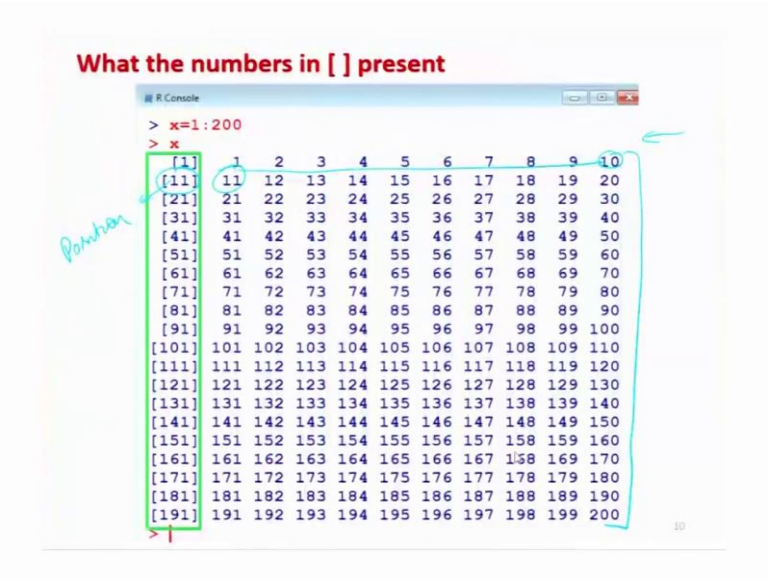
Well, because these numbers are in sequence, so, that is why you are getting the 37 value as the first value in the column also, but yeah, but these numbers which are written here 1 to 200, they can be can be anything actually. But these numbers on the inside the bracket what they are written here they are simply trying to indicate the index of the first value in that line. For example, if you say here 91. What is this 91?

This is the value which is coming out of 1 to 200, but in case if you try to see this 91, which is inside the square brackets this is indicating the location of the first number in this line. Now, there is a confusion. The confusion is this that different people if they try to print the number 1 to 200 on their R console they will get here these number to be different. Why this is happening?

Because actually the width of this GUI window which is here like this, this can be controlled. This is user dependent. Means, if you want to make it smaller or bigger you

can do it. So obviously, in case if this width is reduced for example, then obviously, these numbers are going to be partitioned at 10. So, after this the next value will come here after this for which the 11 number will be there, but this value will be here say 11. So, that is as simple as that.

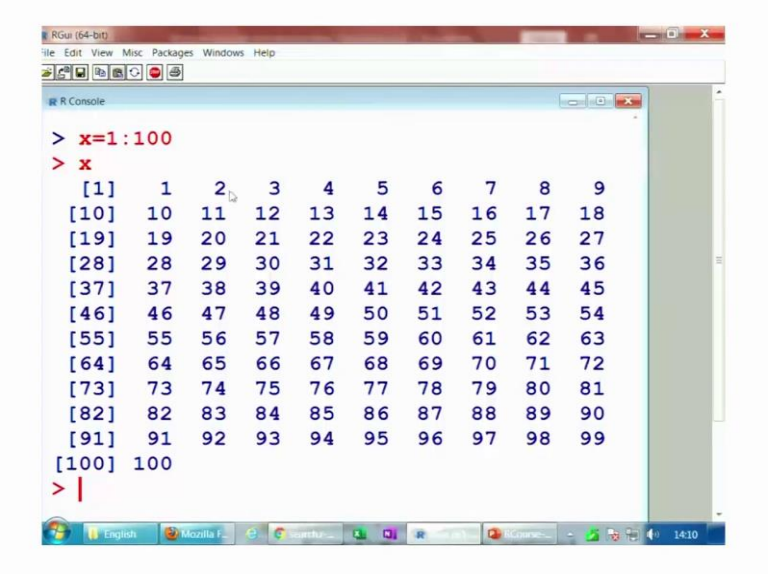
(Refer Slide Time: 18:20)



So, for example, if you try to see here now I have reduced the width of this GUI window and again we have here numbers 1 to 200, but now you can see here 1, 2 up to here 10 and then after that the number at the 11th position is indicated here. So, this is the position and this is the value of the number.

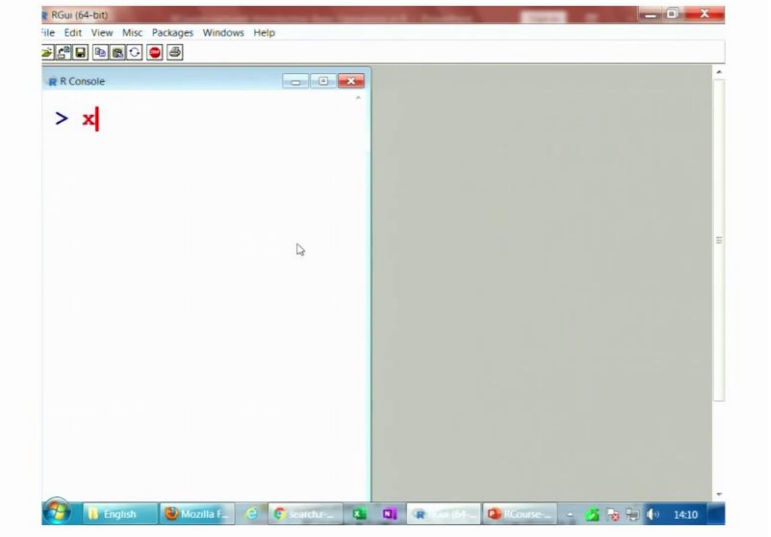
So, let me try to show you this on the R console so that this type of confusion is not there in your mind, when you are trying to work in the R software, right. So, let me try to first make this RGUI window here like this suppose like this.

(Refer Slide Time: 19:03)

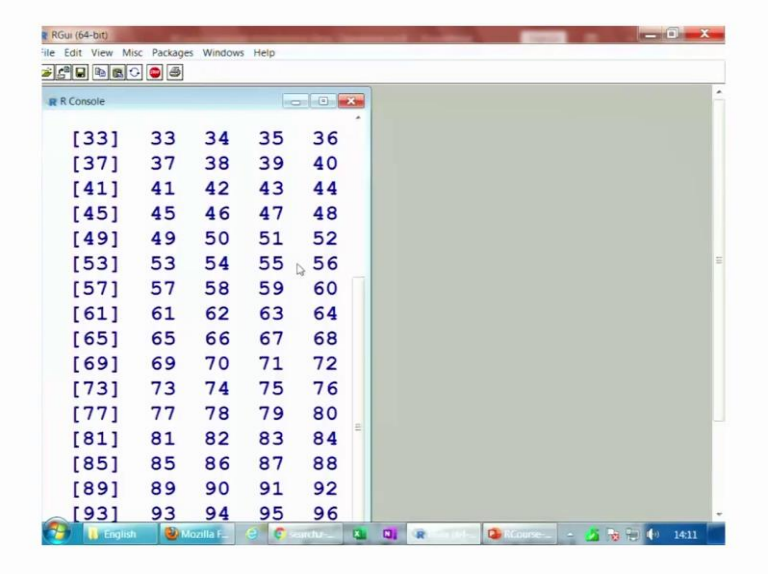


Suppose I try to make it here x is equal to suppose 1 to 100. Now, if you try to see here this here x. You can see here these are the number from here 1, 2, 3 up to 9 and the 10 position is here. And similarly by looking at this 55, I can see the number at the first location, the index of this number is 55. Now, in case if I try to make this window smaller and now in case if I try to see the same thing here again; so, now, I can clear the screen by pressing control.

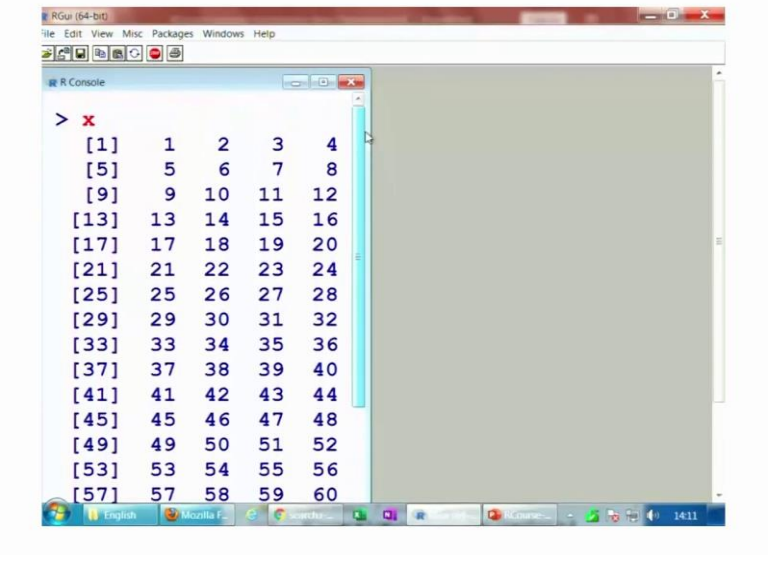
(Refer Slide Time: 19:48)



(Refer Slide Time: 19:48)

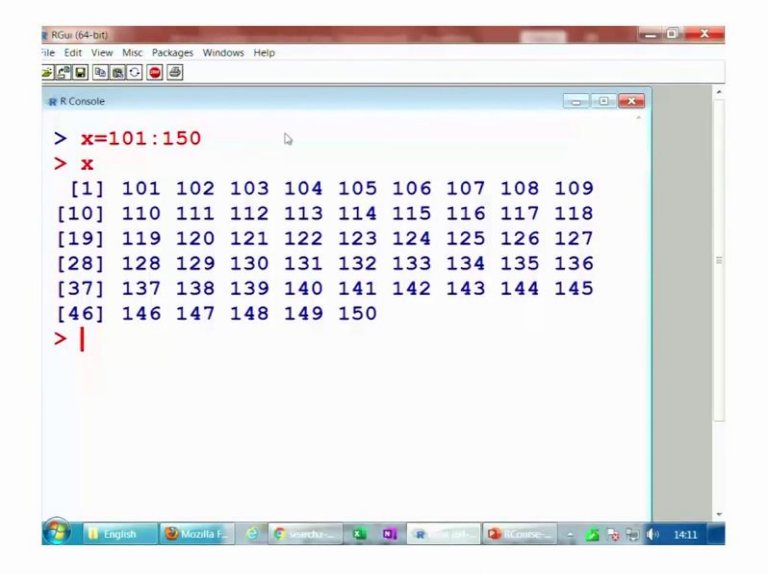


(Refer Slide Time: 19:51)



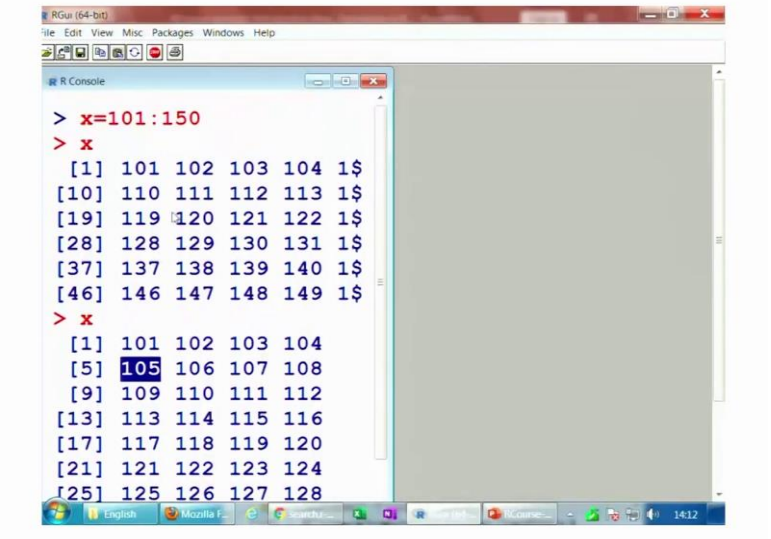
And if I type here x you can see here now there are only four numbers. So, this width is now becoming here 1, 2, 3, 4, 5 in the first row. And then if you try to see this is 5th in the second row and so on right. So, now, ok well this may give you a confusion that why the first number and the index number they are the same.

(Refer Slide Time: 20:11)

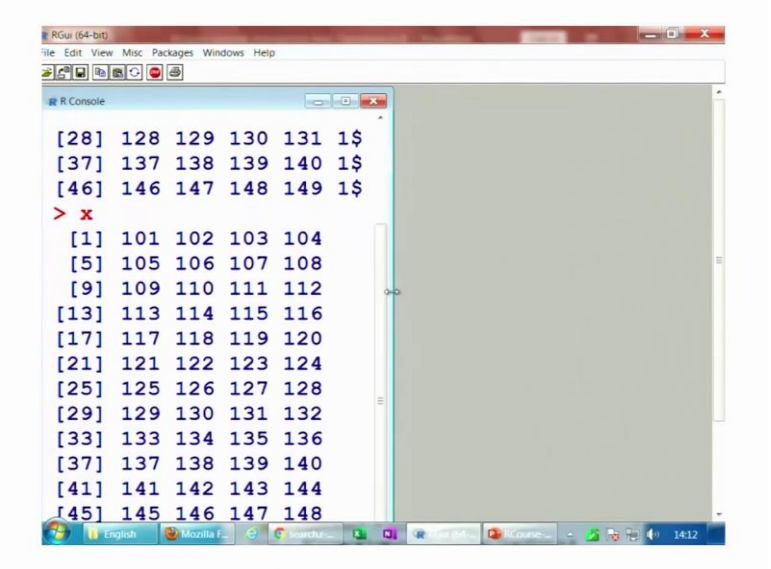


So, let me try to print here some more number. Say 101 to suppose 150, right. So, now, if you try to see here like this; 101, 102 up to here 109, 9 numbers and now that now then we have a position here 10. So, but the number here is 110. So, that means, 110 is occurring at the 10th position. Similarly if you try to see here what is the number at 37th place in the sequence, this is 137. So, this is the type of information is available from here this thing.

(Refer Slide Time: 20:56)

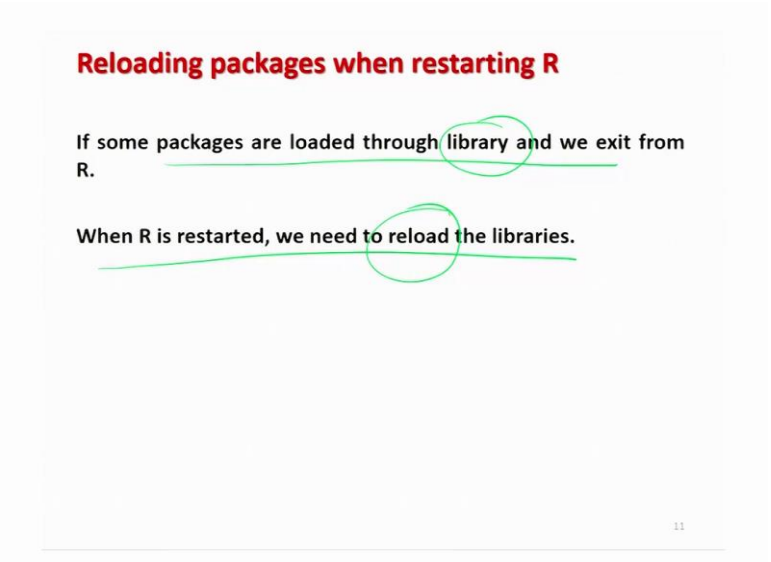


(Refer Slide Time: 20:58)



Now, in case if I try to reduce the width and then if I try to press here x you can see here now this is 101 to up to 102, 104 and after that at the 5th position we have 105. So, you can see here that in the first case the 110 was in the first place in the second row, but now we have here 105 at the 5th place in the second row. So, this is the meaning of this. So, now, I can clear the screen and come back to our slides and I hope you have now understood this aspect.

(Refer Slide Time: 21:28)



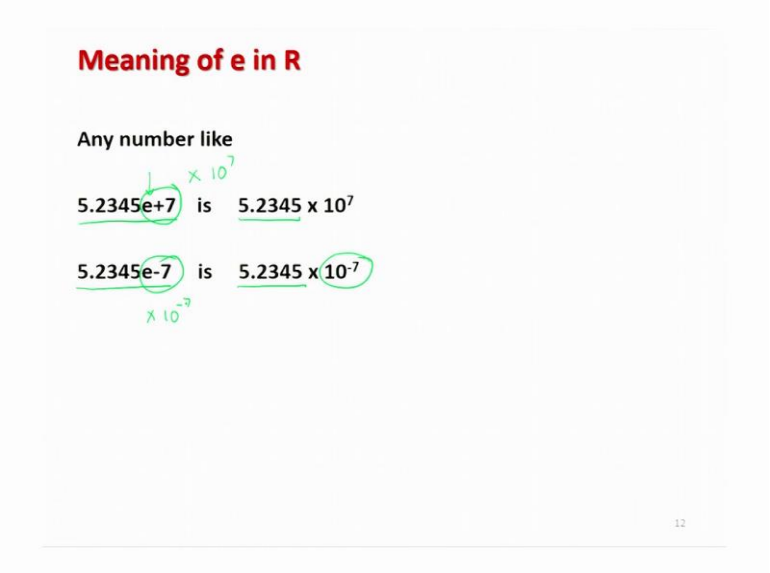
These are very simple thing which many people ask me. So, I thought that ok why not to include it here right. Now, one more cautious that whenever we try to work on the R software then we always install some packages and then working or while using those packages we load them using the function library.

Now, suppose I quit from the R software and I close it. When I try to restart my R software again, then all the libraries in which we were we want to work and which were loaded earlier, they have to be reloaded. You cannot expect that once you have loaded the library this will always be there.

Once you close the R software or you shut down the computer these libraries will be automatically unloaded. So, many times it has happened that you are working on your computer and something happens and then you have to restart your R software and after that you start working, but it gives you means some error some problems and the people are unable to understand what is happening.

Because they are thinking ok, just now we had loaded this library, but since you have restarted the computer or the software those libraries are unloaded and you need to reload them again ok. That is a very simple ok.

(Refer Slide Time: 22:59)

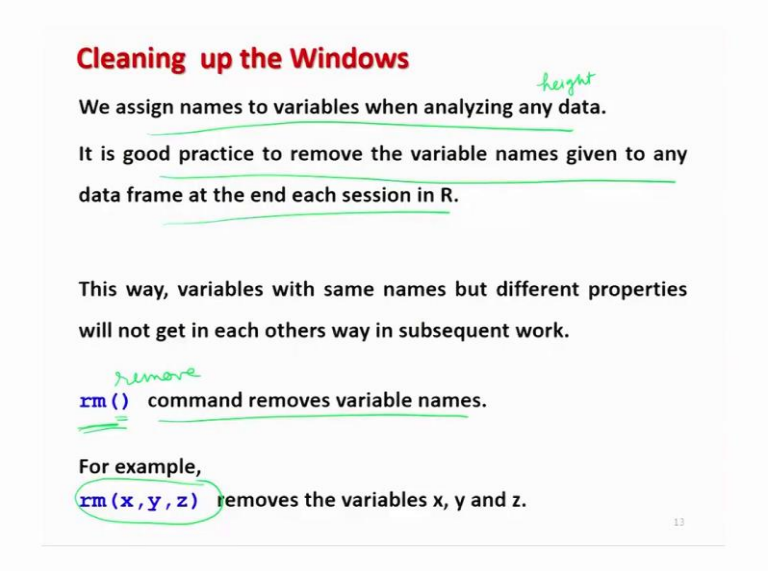


Another point of concern. In R, many times you will see the numbers like this. For example, I am writing here 5.2345 then after that I am writing here an alphabet is small e

and then plus 7. What is the meaning of this? This number has a very simple meaning. It is 5.2345 into 10 raise to the power of 7.

So, this part this is something like multiplied by 10 raise to the power of 7 and similarly in case if it is written like this, 5.2345 e to power of minus 7, this means this number is 5.2345 into 10 raise to power of -7. So, this e minus 7 is like multiplied by 10 to the power of minus 7. Yeah. Sometime people try to misinterpret it like as exponential function or something else. Please do not do this thing, ok.

(Refer Slide Time: 24:04)



Now, when we are trying to work in the R software, then we need to clean up the windows also. And sometime we want to remove the variable. For example, when we are working or doing some programming in the R software we always assign some names to the variable, right.

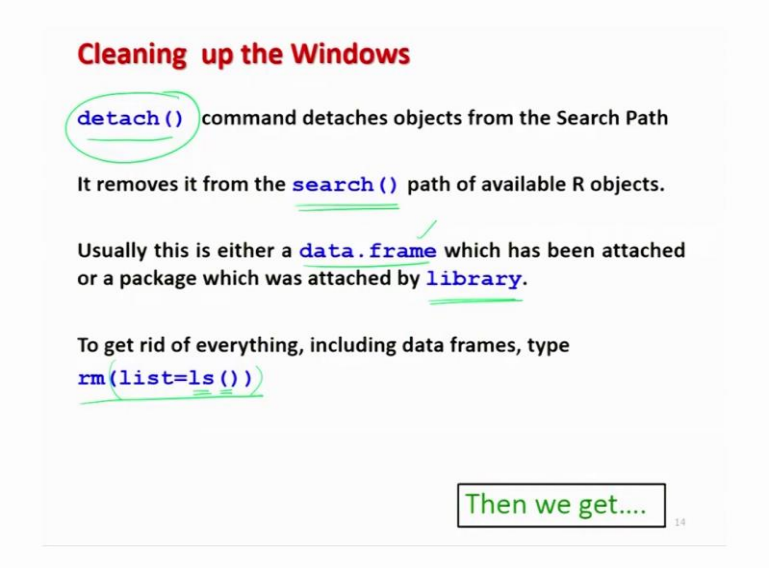
For example, if there is some data on the height of the student. So, I can define here a variable like here height and I will store the data in this variable. So, now, this variable will remain stored in the R software directory as long as you do not remove it yourself. And now what are the consequences? The consequences is that the consequences are that suppose you start working on another problem and you defined the variable once again as height.

Now, what will happen? That all the values which you have stored or which you had stored in the variable height in your earlier problem, this new variable height and its value will overwrite it. The older values will be removed and the new values will be there. And for that, the R does not give you any option also. R will not ask you whether you want to overwrite or not. Once you have pressed the enter after the things are gone.

So, in case if you are trying to work sometime you want to clean up the window you want to remove the variables and actually it is a good practice to remove the variable names given to any data frame or any data set at the end, after you finish your work in the R software. So, this way what will happen? This type of confusions will not be there and yeah and then there would not be any problem. So, the question is how can you remove the variables.

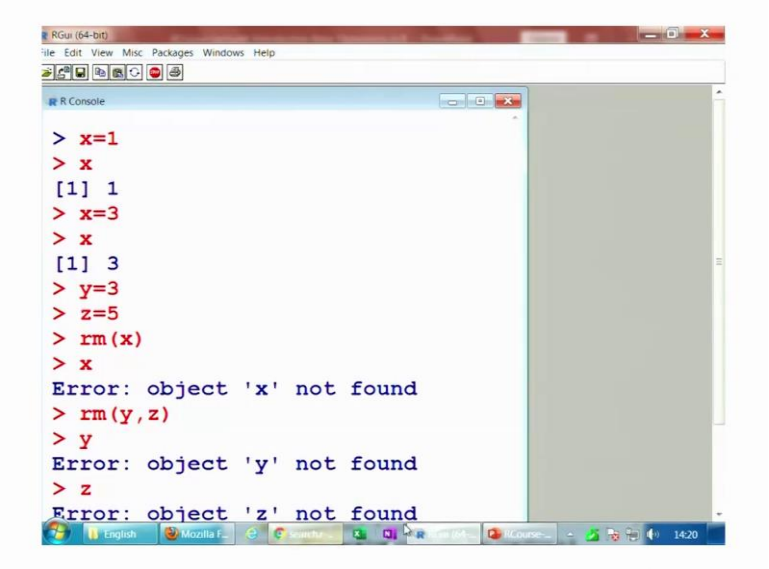
So, we have a command here rm. rm means remove, it is the short form right. So, if you type here r m and inside the parenthesis you write the variable name, this command will remove the variables and you remove one variable or more than one variables at the same time. For example, if you want to remove three variables say x, y and z then you can write down here rm inside the parenthesis x comma y comma z and this will remove all the three variables from your R software from your working directory.

(Refer Slide Time: 26:57)



And before I go further let me try to show you here these things right.

(Refer Slide Time: 27:03)



So, suppose I try to make it here, I take here one value here x equal to 1 right. Now, you can see here x is equal to here 1. And suppose by mistake, I write down here x equal to 3. Now, as soon as I press here enter it will not ask me that there is already an x which was present, but now if you try to see the value of here x this is overwritten and now it is only 3. And similarly if you try to say here y equal to 3, z is equal to here 5 and so on.

Suppose I want to remove this x, y and z. So, first option is this I can see here rm and inside the parenthesis I write only one x. So, now, you can see here x if you try to see what is this? This is removed and it is giving you a message that error object x not found. That is obvious because you have removed this variable. And similarly if you want to remove here y and z together then you can write rm inside the parenthesis y, z and now if you try to see here why this is missing z, this is missing right.

So, this is how you can remove these variables when you are trying to work. And similarly if you recall we also had discussed one command detach. So, this was the command which removes the object from the search path, right. And it removes from this from the path of this parenthesis command which are available for the R objects. Actually this detach command is a very general command.

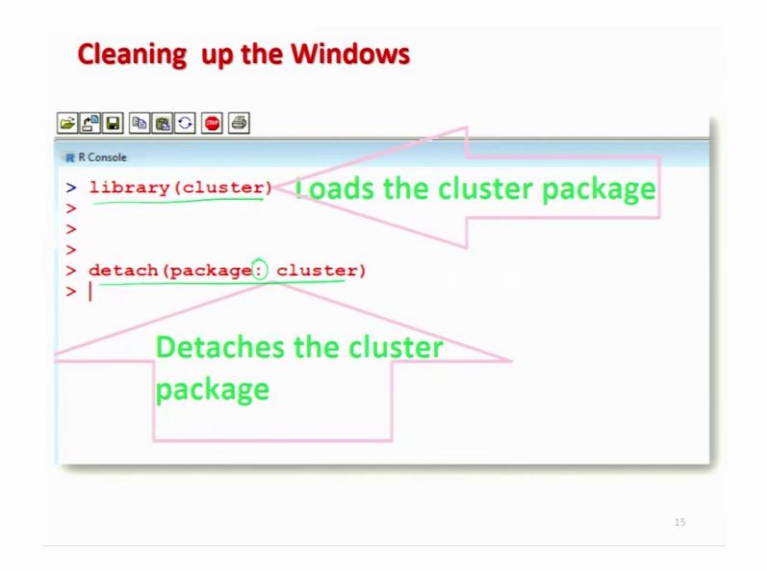
This is used for functions and data frame and other type of thing. Although we have not done it here up to this point, the concept of data frame and other things, but at this

moment, I would like to inform you and later on when we are trying to work on this concept explain you once again, right.

So, all those data frame, data sets function etc. which have been packages which have been attached with the library function they can be removed by this detach function. And above, all in case if you want to remove everything including the data frame also along with the variables, you can use here the command rm list is equal to ls and then this parenthesis and you write this inside the parenthesis.

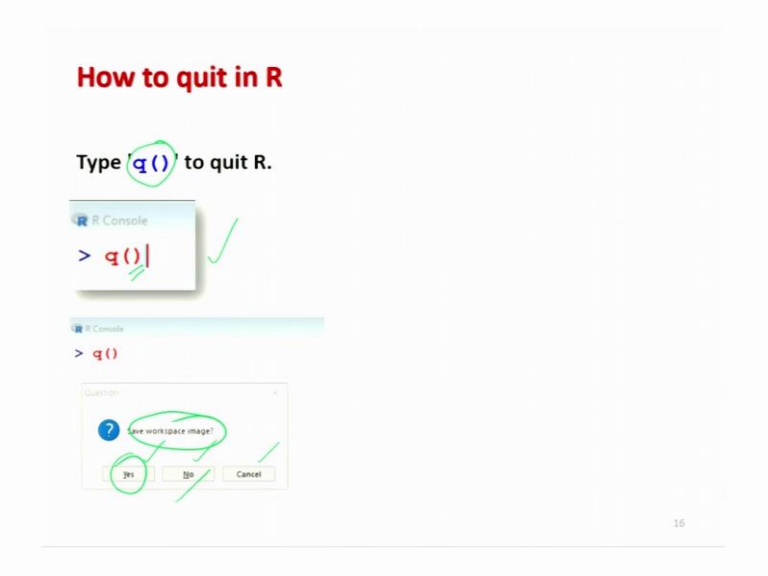
I will not execute it here on my computer otherwise everything will be cleaned. But I am telling you that this is the command to remove everything from your directory including the data frames also.

(Refer Slide Time: 29:50)



So, for example, I can show you here, suppose if I try to upload a directory using the command library cluster. So, this will load the cluster package and then if I try to write down here detach package cluster means package and cluster then it will remove the package from the library from the search path. And then in case if you try to use it again it will say that ok this is not available.

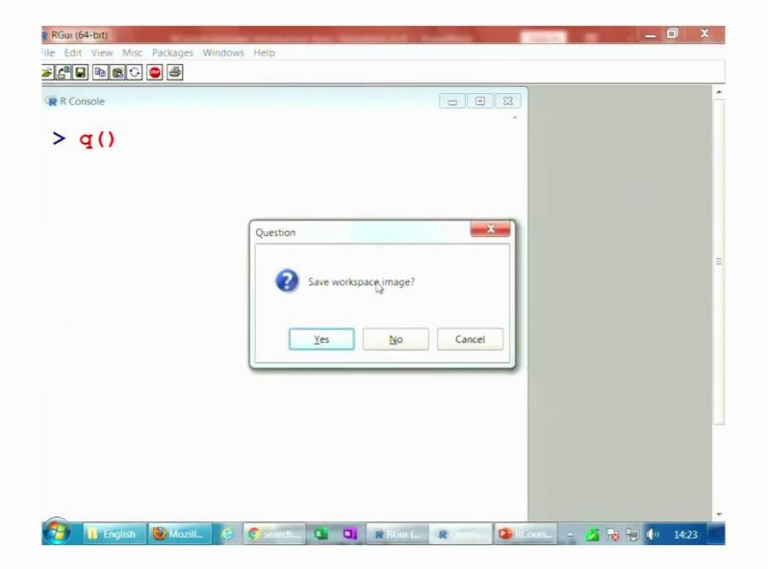
(Refer Slide Time: 30:19)



Now, finally, when you have completed your job and you want to quit the R then the command here is q and then parenthesis. So, once you try to write down here q in this parenthesis on the R console, this will ask you do you want to save the work space Yes or No or Cancel. So, what is the meaning of this work space image?

That means, whatever I have done up to now, whatever variables we have defined in this session, whatever data set we have imported in this session, they will all be saved otherwise everything will be lost in case if you press here Yes. And if you do not want to save them simply say No that is the very simple thing. So, let me try to show you that what happens with this here quit on the R console itself.

(Refer Slide Time: 31:22)



So, you can see here I can now clear the screen by pressing control L and if I try to make it here quit, as soon as I say enter, so, it is asking me, save work space image, Yes, No. Now if I say Yes or No, accordingly it will work, right, ok. So, now, we come to an end to this lecture.

So, you can see that in this lecture that was a very basic elementary lecture in which I have covered a number of small topics which are very important for you to learn when you are working in the R software. Well, there are some more such commands which are there and I will try to cover them in the next lecture. This otherwise; that means, too many commands will make you more confused.

So, that is why I have taken this collected number of commands. Well, some of the commands I already had taken in the couple of last lectures. Well, they were the need of the time and the reason is this means I can share with you very honestly that whenever we are trying to learn the software, we have two options.

The first option is that first we try to go through with the theories and then after that we try to come back to the software and try to learn the commands or the second option is this we simply jump into the well. That means we simply start the software and we start doing something.

And in this process we will learn something at that moment itself and there will be a few things that we will learn later on and for that only I always say, ok, we will try to learn it in the forthcoming lectures. But, I personally believe that this second approach makes the learning more interesting and it is easier for you to understand. Yes, I agree that in the beginning there will be many commands which you have not seen.

But I have taken care that I am not trying to take any complicated commands, but I am trying to take very simple commands which are not so difficult to understand. And as we are moving further gradually you will see that you will learn all those commands and at the end, there would not be no question like we will try to see it in the next couple of lectures.

So, you try to have a look on these commands what we have learned today, try to practice them, try to revise them and I will see you in the next lecture. till then goodbye.