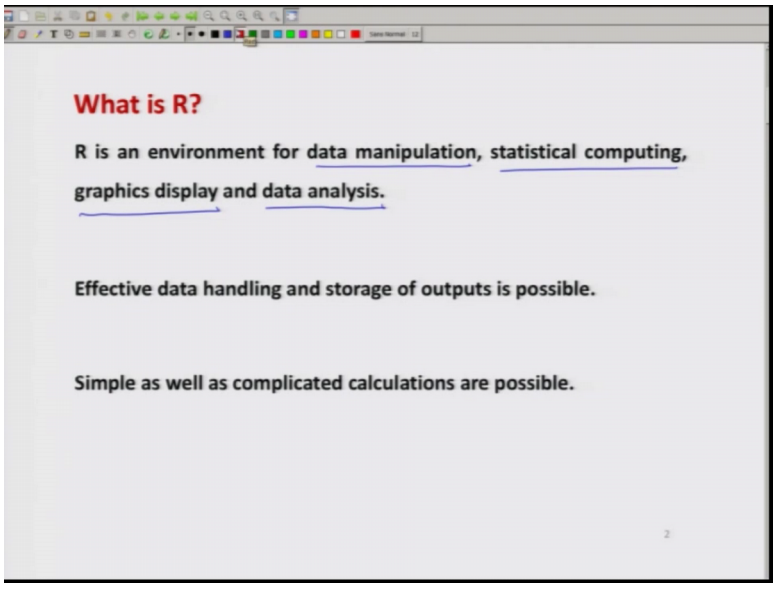
Introduction to R Software
Prof. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture – 01
Why R and Installation Procedure

Welcome to the course on Introduction to R Software. Well in this course, we are going to learn about different aspects of software what is called as R. So, in this lecture, we are going to understand that why should we learn R and how should we install the R software and related software on our computer the basic idea is that I am going to start this course at an very elementary level. So, that anyone who has no idea about even about how to install the software or how to operate the computer even he should be able to understand it. So, let us try to start the course with this first lecture.

So, first question comes what is R? R is essentially an environment for the data manipulation is statistical computing as well as graphical display and data analysis, right. R is just like any other software there are different types of software which helps us in mathematical calculation and statistical data analysis similar to them, R is another software right and R has an advantage that R can do data manipulation, R can do is statistical computing as well as simulations, R is capable of graphical display and R can also help us in doing different type of data analysis.

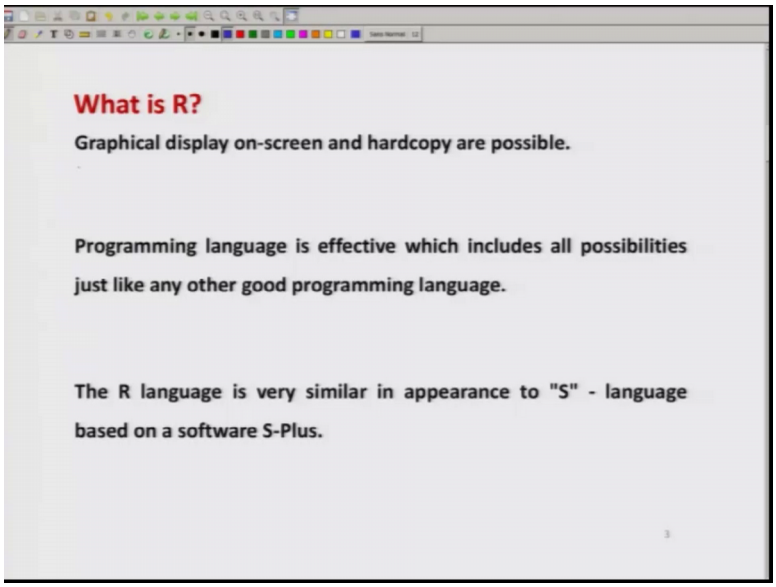
(Refer Slide Time: 01:59)



And just like any other good software, R also has an effective way of data handling. So, it can handle the data easily and it can store the input and output variables. This can store the outcome in the form of a scalar as well as form of a vectors or a matrix, right and in R software, simple calculations are possible as well as complicated collections are also possible and it is not a difficult the mathematical calculation like addition subtraction and this vectors and matrices everything is possible just like any other good software.

Whenever we are doing any software handling, there is also a requirement of the graphical output. So, the R is capable of graphical display also and these displays can be made over the screen as well as we can save them in the postscript file, PDF, file jpg file or say any other type of file from which we can obtain the hard copies of the graphics right.

(Refer Slide Time: 02:55)

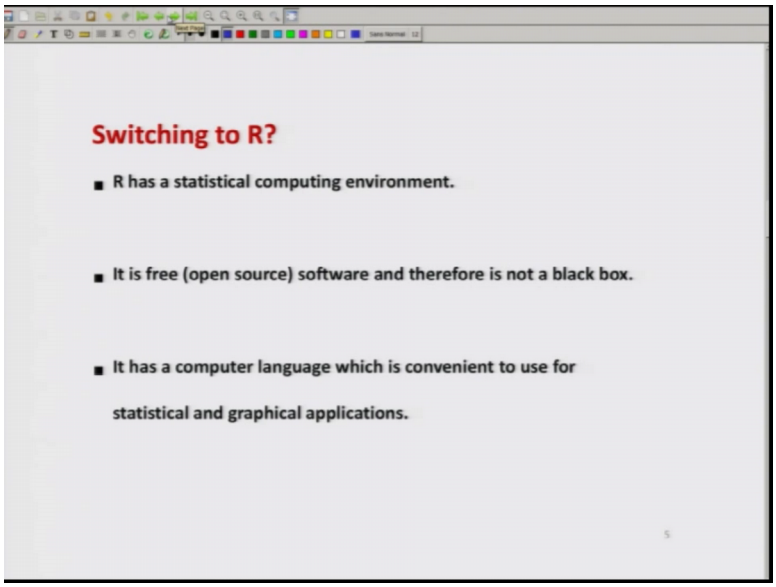


And R software has its own programming language for example, we have learned different types of programming language like as Fortran, C, C++ and so on. Similar to that R also has its own programming language, right and this programming language is very effective and it includes all sorts of possibilities just like any other good software programming language, right.

The programming language of R is very very similar to another programming language what we call as S language actually earlier there was a software what is called as S plus

and this software S plus had a language what was called as S language later on the R software was developed on the lines of S plus software and the programming language of R is very very similar to the programming language of S plus.

(Refer Slide Time: 04:32)



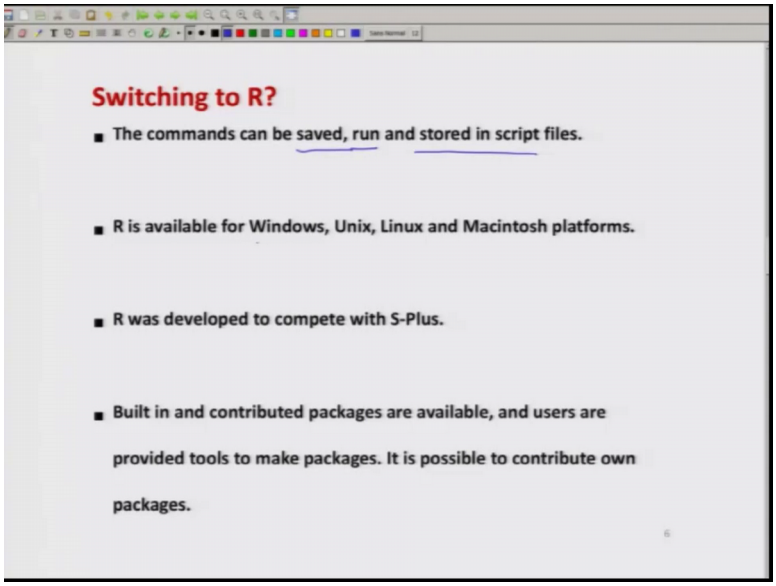
Now the next question comes over here; when there are so many software, then why should I use R or why should I switch to R software. So, it is kind of situ to understand these facts well the first advantage is that R has an R has a statistical computing environment R can do all sorts of computation; what any other software can do.

The biggest advantage of R is that this is a free software there is 0 cost and R has an open source what do you mean by open source for example, in case if I try to use any software then we really do not know in most of the cases that what is happening inside the program usually we simply give the input and we obtain the output what is happening in between that is usually unknown. But this is not true with R software right it is not like a black box, but it is an open source this is uploaded on the website of R and anyone can look into the program that how the computations are being made.

So, in this sense, this is an open source the language of R programming is also very very convenient and anyone can do a statistical and graphical applications using this R programming another advantage of R is that whenever we are trying to execute a command, then those commands can be saved and those commands can be a run and

moreover, they can be stored in a script file what does this mean that whenever we are trying to execute a program first we have to write down the program in a file and then we have to use further execution. So, all these commands they can be saved in a separate file. So, that if you want to retrieve it later on you can easily do it.

(Refer Slide Time: 06:56)



Another advantage of R is that this R software is available for all the platform this is available for windows platform Unix platform, Linux platform as well as Macintosh platform and actually as I said earlier this R was developed essentially to compete with S plus S plus is a very good software and, but it is a paid software. So, one has to pay the cost to buy it. So, people started developing this software to compete with S plus and in this software there are 2 types of packages one type of packages to execute a particular task they are built in inside the software and another type of packages are which are actually contributed packages what do you mean by contributed packages that suppose I am doing a research or I develop a new tool and suppose I write the program to execute the new tool in the R programming language, I can submit this program to the developers of R and they will try to verify it, they will try to check it and if they find the cut everything is alright then they can upload my program over the site of R.

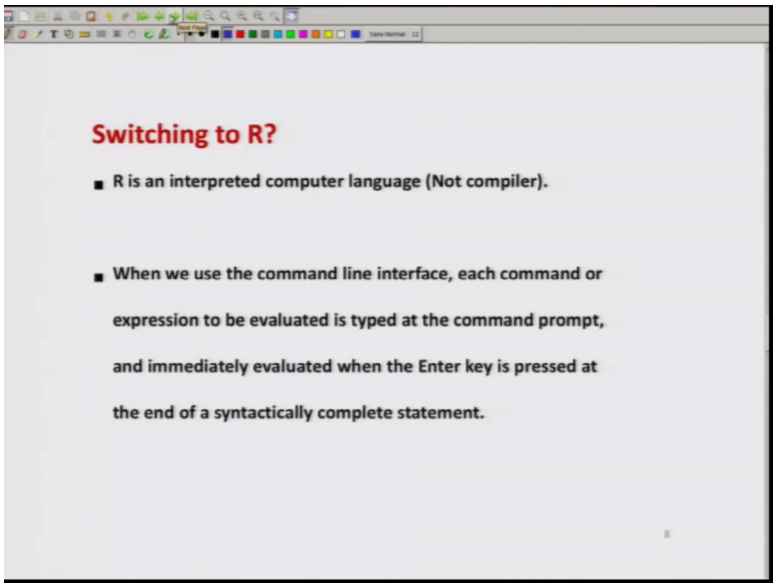
The advantage is that that anyone who wants to use my develop tool he can simply download the program and can execute its data analysis. So, in this R, we have both types of packages one built in packages as well as contributed packages and the users are

also given an opportunity to make their own packages that is another advantage of here R just like any other language; the R language also has all sorts of the structure this can provide the logical control of branching clickers can provide the looping as well as this can provide the modular programming using the concept of here functions; what does this mean.

For example, any software language has several options we can do the mathematical manipulations we can do the logical manipulations and sometimes, we have to repeat the program for that we try to use the concept of loop and in many many cases whenever we are trying to deal with a huge program then the entire program is divided into several smaller parts and every developer develops a smaller part and then they are combined together. So, that is called as a modular programming.

So, all these things are possible in R one can do the mathematical operation logical operation one can provide the loops for the repetition of the program as well as one can do the modular programming when we try to execute any program on R then there are some error messages, if there is some problem in the programming and these error messages are helpful in identifying the problems in the program and to execute the program. So, just like any other software R also provides us the error messages which helps us in the programming and execution of the program.

(Refer Slide Time: 10:20)



Next aspect of R is that that R has an interpreter and it is not actually a compiler what is the difference between interpreter and compiler whenever we are writing a program; program is simply a set of some commands which are executed together. So, suppose we have written 5 commands in a file and now when I try to execute it there are 2 options first option is that the execution will start from line number 1 line and then it comes to line number 2, then it comes to line number three and so on and in every line it will try to identify and it will try to check if everything is alright or not is there any mistake in writing the program.

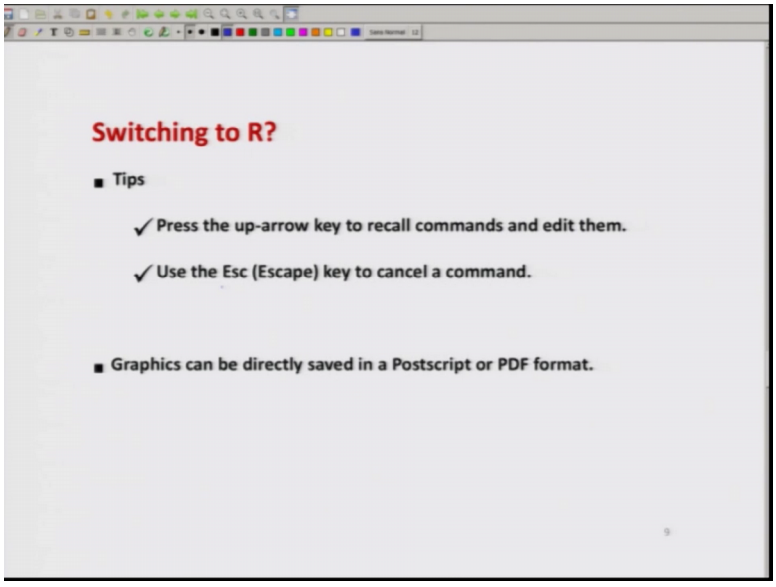
So, in interpreter what happens that when the control starts from line number one then it comes to line number 2 now suppose there is a problem in the line number one then the program we will stop there and it will give us an error message and unless and until we look into the error and we try to correct the program the control will not go on the second line and similarly, when I try to correct my first line then the control comes on the second line suppose second line has no issues. So, the control will come on the third line. Now suppose on the third line there is some error there is some problem in writing the program.

So, now the program will stop there and it will ask us to first correct the error on the third line and only after that it will go to the fourth line whereas, another option is to use a compiler in compiler what will happen the execution of the program will start and it will start from line number 1 line number 2 and line number 3 and it will reach to the last line and in every line, it will try to find out the problems and errors and then it will try to compile all the errors in a separate file and then the program will be executed. So, after the execution of the program we have to look into the outcome and then we have to see that which lines have errors and then we have to correct the entire program.

So, in the interpreter the errors are given at every line at every stage whereas, in the compiler the errors are given only say all together. So, R has an interpreter that it will go line by line and as soon as it finds a mistake in a line that will stop the program over there and first you have to correct the mistake and then it will move to the next line right another thing is that there that whenever we are trying to use the R then as soon as we try to type the command over the so called command line do not for this language means very soon, I will try to show you how the software looks and how these command line interface and all these technical language they are they are understood, right.

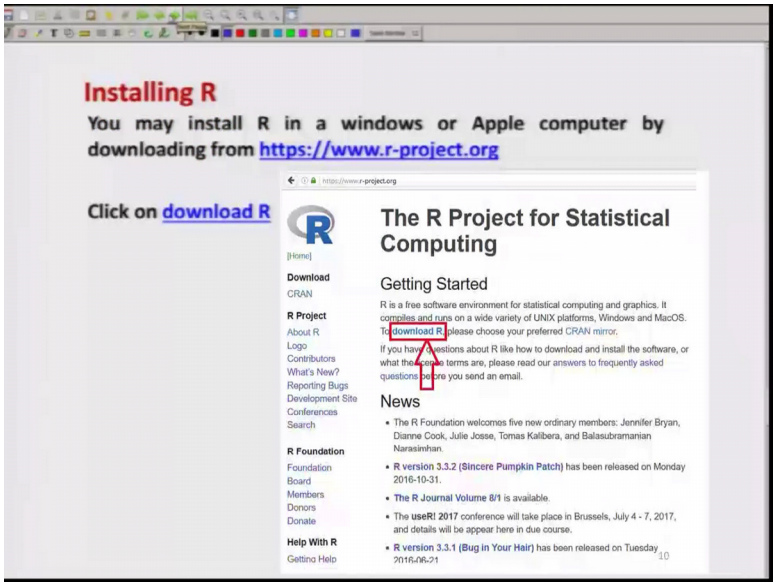
So, as soon as we type a command and we press the enter key on my keyboard, the program is executed, right. So, that has a very simple thing that if you do not want to write the entire program, but if you want to make only a one line statement you simply have to type it and then press the enter key on your keyboard and the program will be executed alright. Now some tips that whenever you are trying to play with the R software you can use the arrow keys on your keyboard to move to the earlier commands and this will help us in editing them; suppose you are writing something and you want to cancel the command then in that case you can simply use the escape key that there is a key on the computer which is something like Esc.

(Refer Slide Time: 14:19)



So, you simply have to press the escape key and the execution of the command will be canceled and as I said earlier that in case if you want to do with the graphics different types of plug different types of graphs different types of graphics can be constructed using the R software and they can be saved in a postscript file PDF file as well as they can be saved in a I say this jpeg format PNG format they can simply be copied and pasted into another file and as go on, right.

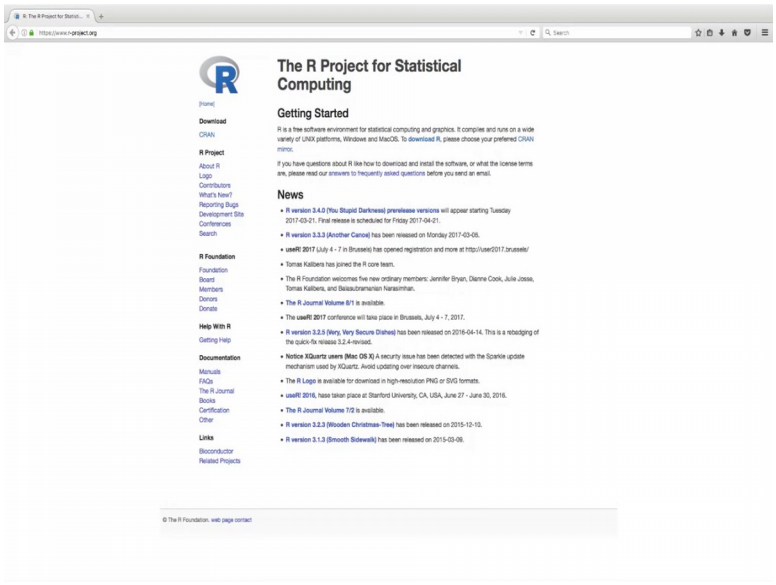
(Refer Slide Time: 14:58)



The next thing which comes that how should we get R and how to install it on the computer, right. So, now, let us try to understand this thing this is a very simple thing means if you have a little bit idea about using the internet anybody can do it right. So, in order to install the R; what we might try to do there; there is a website this website is www R hyphen project dot org right and what we try to do here that we try to execute this command. And then I will try to show you what are we going to get and the same outcome I am trying to write down here if you try to see I have simply copied and pasted, but you, but in order to make you more confident; what I will try to do here that that I will try to show you this online, right.

So, www R project dot org if we try to see we have got this website right.

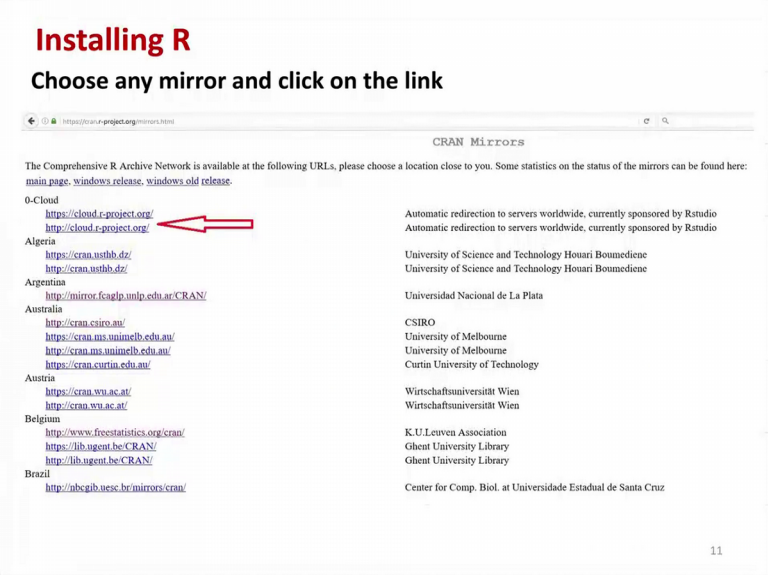
(Refer Slide Time: 16:11)



So, you can see here now what I have done just for the sake of convenience and in order to illustrate it what I have taken a screenshot of this webpage and I have copied it here. So, I am showed that this will not create any confusion for you, but it will help us in understanding that things. So, once you come to this homepage then what you have to do here that you need to go here, there is a command here download R. For example, I can show you here for example, you can see here there is a download R here and here you double click it and once you double click it, it will give you this home page right that is the same thing which I am trying to show you here also.

So, you can click over here at the download icon and then in the next site, you will get here this type of site what you have obtained.

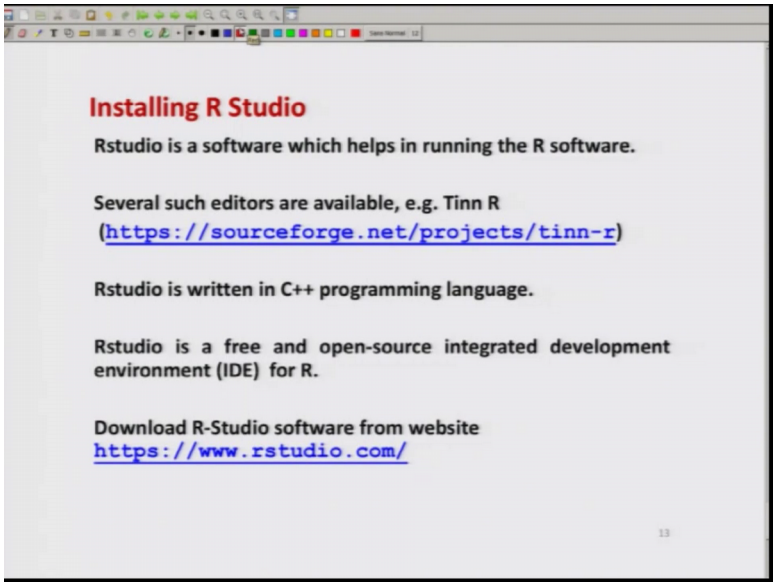
(Refer Slide Time: 17:17)



So, on the Philips hand side you can see here that there are different types of addresses which are given. So, actually different people in different countries they have uploaded the software. So, you can click on any of the link and it will open the page for downloading the R software. So, you can just click over here or say here whatever you want right and then once you do it, then you will the software over here for example, I can show you here that once you try to do it here. For example, if I try to see here in Austria, suppose I try to bring over here it will just go to that place and it will download the software right you can see over here and here you can see that you can here download the thing.

So, let us now come back to our original slide where I have just taken a screenshot and I have pasted it here for a better understanding. So, now, you can see here and that there are here different types of link one link is download R for Linux this is for Macintosh and this is for here windows. So, depending on the platform which you have you can just double click on it you can download the software and after that you have to simply double click on the software and then the and the software will be installed on your computer right after this, I will just you to install another software also.

(Refer Slide Time: 18:54)



Now, after this we come to another aspect whenever we are working with R software we have 2 options, either we can work directly with the R software we can execute the command inside the R software and second option is that I can take help of an of a supplementary software which works inside the R from outside. So, so there are different types of software which are which are available and in this course I am going to use a software what is called as R studio right. So, R studio is essentially a software which helps in the execution of R software and beside R studio there are other types of software which are available for example, one of the another software is the Tinn R and this can be downloaded from this site and, but I can use any one actually I am not trying to say at all that either Tinn R is better than R studio or vice versa means I have chosen R studio to work. So, this R studio is actually written in the C plus plus programming language.

R studio is also a free software because in open source software and this can be freely downloaded from this website. So, what we have to do here that we simply need to copy this link and then we have to open it in the internet browser. So, for example, I can show you here that if I try to say here type here R, r studio dot com then I get here this thing. So, I have simply taken here of a screenshot of this thing. So, what we need to do this software will be opened and now we have to simply come over here and then I have to click over here at the download part and this software will be downloaded and once this software is a downloaded then you can install it on the computer. So, essentially you will see that once you have installed the R software and R studio software you will have a

link like this here R and clicks here R studio and now we are ready to move into the learning of R software.

So, this was the first lecture in which I have tried to give you a basic idea that why should we use R and how should we install the software on our computer, how to obtain the software and possibly this will help you in getting ready for learning the course content of the R software in the next lecture. So, we will see you in the next lecture till then, good bye.