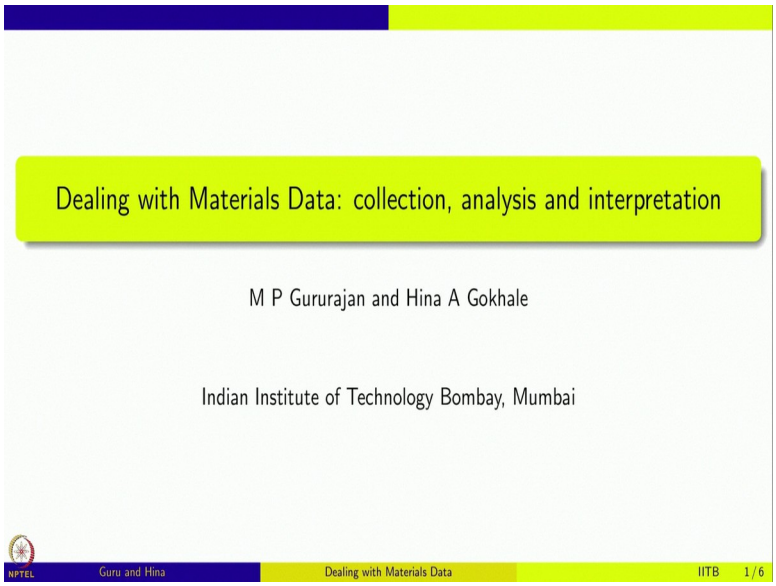
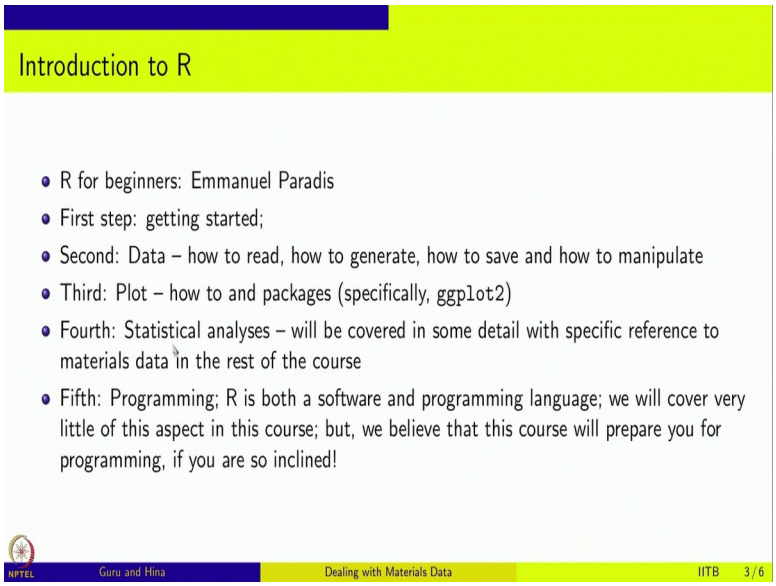
Dealing with Materials Data: Collection, Analysis and Interpretation
Professor M P Gururajan
Professor Hina A Gokhale
Department of Metallurgical Engineering and Materials Science
Indian Institute of Technology, Bombay
Lecture No. 18
Introduction to R: Summary of the Module

(Refer Slide Time: 0:29)



Welcome to dealing with materials data. In this course, we are going to learn about collecting, analysing and interpreting materials data. This is the first module. This is an introduction to R module. So, in this module we are learning how to use R for data analysis and interpretation, some preliminaries of the R programming language, we are going to learn more about the data analysis and interpretation parts as we move along in this course. So, this is the last session of this module. So, this is a summary session.

(Refer Slide Time: 0:57)

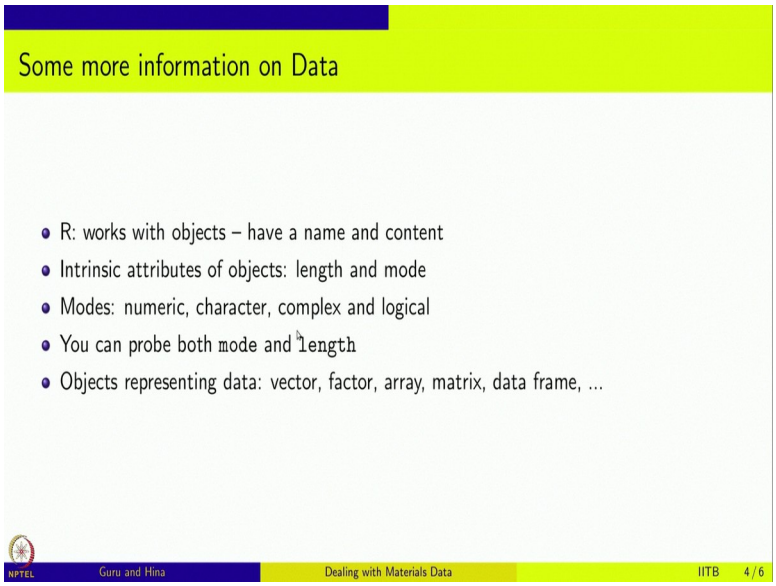


And so, if you look at a book like R for beginners, by Emmanuel Paradis. There are about five sections in the book. The first one is getting started. The second is called data, how to read data, how to generate data, how to save data and how to manipulate data. Third one is plotting, how to isolate some packages. For example, we have used ggplot2.

4th is statistical analysis, and that will be covered in great detail with specific reference to materials data in the rest of this course. And the fifth is related to programming. So, R is both the software and the programming language. And like I said, we are giving just a tutorial introduction to R, we are not going to cover lots of programming aspects. But I do believe that this course will prepare you for programming and if you are so inclined, then this should be a nice starting point for you to start.

So, if you look at the material that we have gone through. Of course, we have learned how to get started with R, we have also seen how to read or generate and save or manipulate data. And we have looked at plotting in little bit of a detail. And statistical analysis is something that we have not started. But very next module, we are going to start, for example, is on descriptive data analysis.

(Refer Slide Time: 2:35)

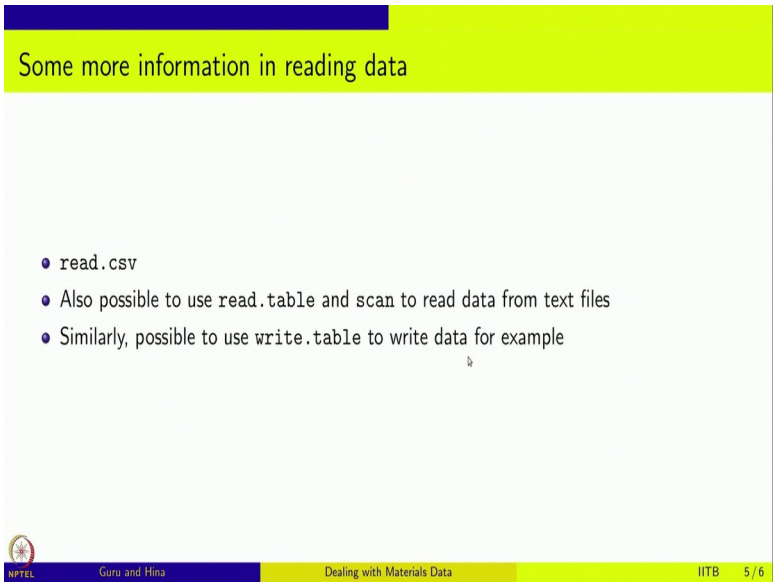


And, so, but there are some more aspects that has been left out. And so I am just going to mention them here. And I recommend that you read more about this with the available material online. So, R works with objects as you have noticed, and each of these objects have a name and content. For example, when we have this variable name element or X and it has some content. So, it is 15 observations and 4 variables, etc.

And the objects have intrinsic attributes like the length and mode. So, if you, you can ask for length of an object and it will tell you what is it size and the mode is like whether it is numeric or character or complex or it is a logical like true, false, etc. So, and of course, you can use both mode and the length commands to probe the intrinsic attributes of any given object.

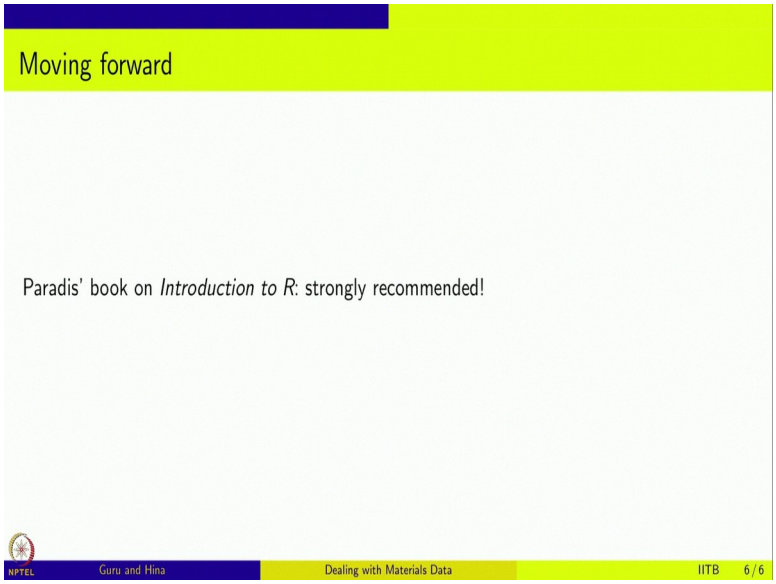
And objects representing data can be of many different types of vector, factor, array, matrix, data frame, etc. We have seen tables and matrices and data frames, etc. So, so there could be much more detailed information and this might become very important if you get into R programming.

(Refer Slide Time: 3:51)



And reading data we have used read dot CSV, but it is also possible to use read dot table and scan to read data from text files. So, this might become important in some cases. So, so you can learn about them. And similarly, you can use write dot table to write data for example.

(Refer Slide Time: 4:12)



So, to move forward, we are not going to spend too much of time doing this other aspects. So, this introduction to R is a very, very tutorial introduction. So, it is more focused towards what we want to do for the rest of the course. So, sometimes all the details are not given. And also when you are programming, there is not just one way of doing things, there could be more than one way of doing things.

And there could be gradation among these different ways of doing things, maybe some or more efficient, computationally or more optimal than others. So, those aspects we have really not looked into. So, sometimes we have done very circuitous things, maybe, maybe because that is just the one way of doing it or just straightforward way of doing it and maybe there are better ways of doing it.

So, so, all these aspects we have not discussed and however there are several references that I have given so far and I want to add that this Paradis book introduction to R also. Because it is a very short book running into less than hundred pages, but it discusses all these five aspects like I mentioned earlier, and so, it can add to what you have already learned and it can prepare you for programming an R if you are so inclined. But we will continue with our tutorial mode and we will start using R for doing some statistics and data analysis for the rest of this course. So, thank you.