Learning Analytics Tools

Professor Ramkumar Rajendran

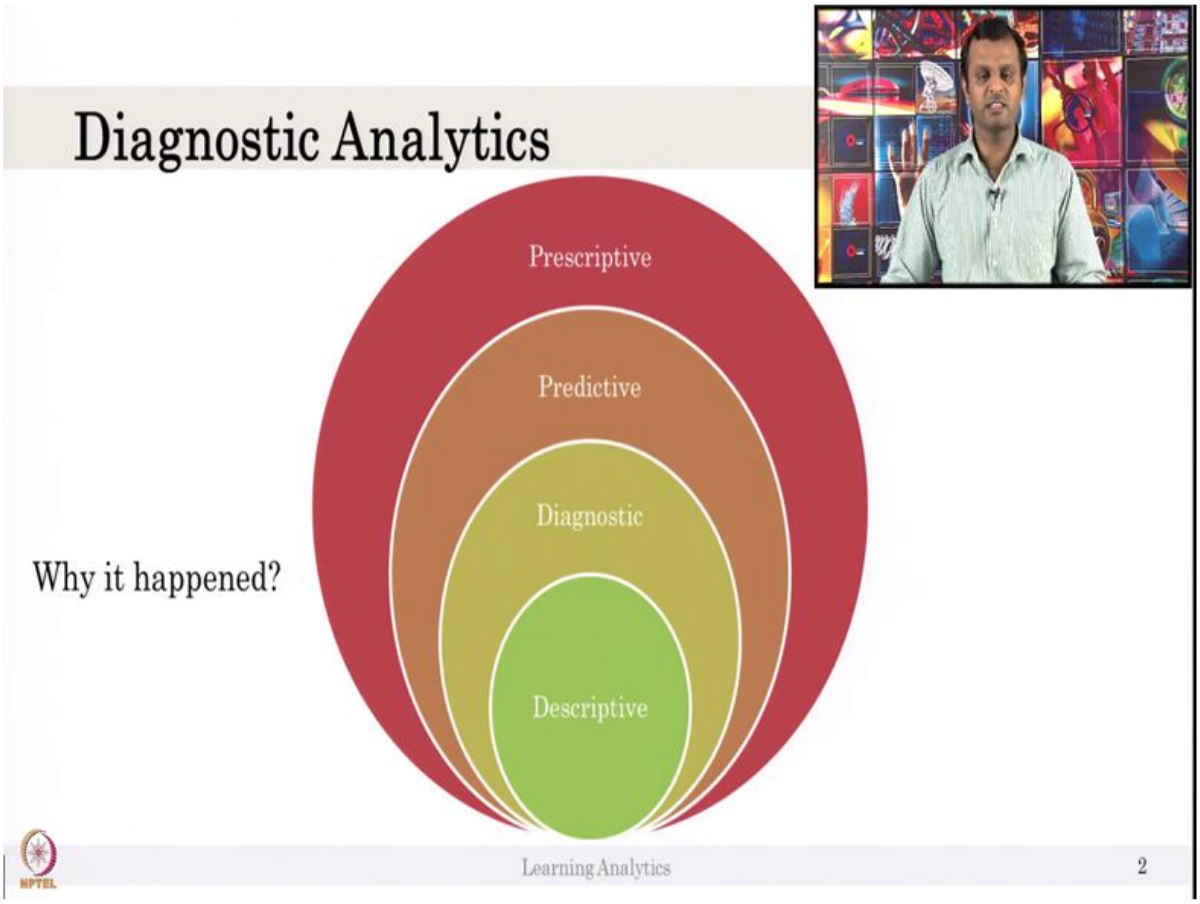
Educational Technology

Indian Institute of Technology, Bombay

Lecture No. 5.1

Diagnostics Analytics

(Refer Slide Time: 0:24)

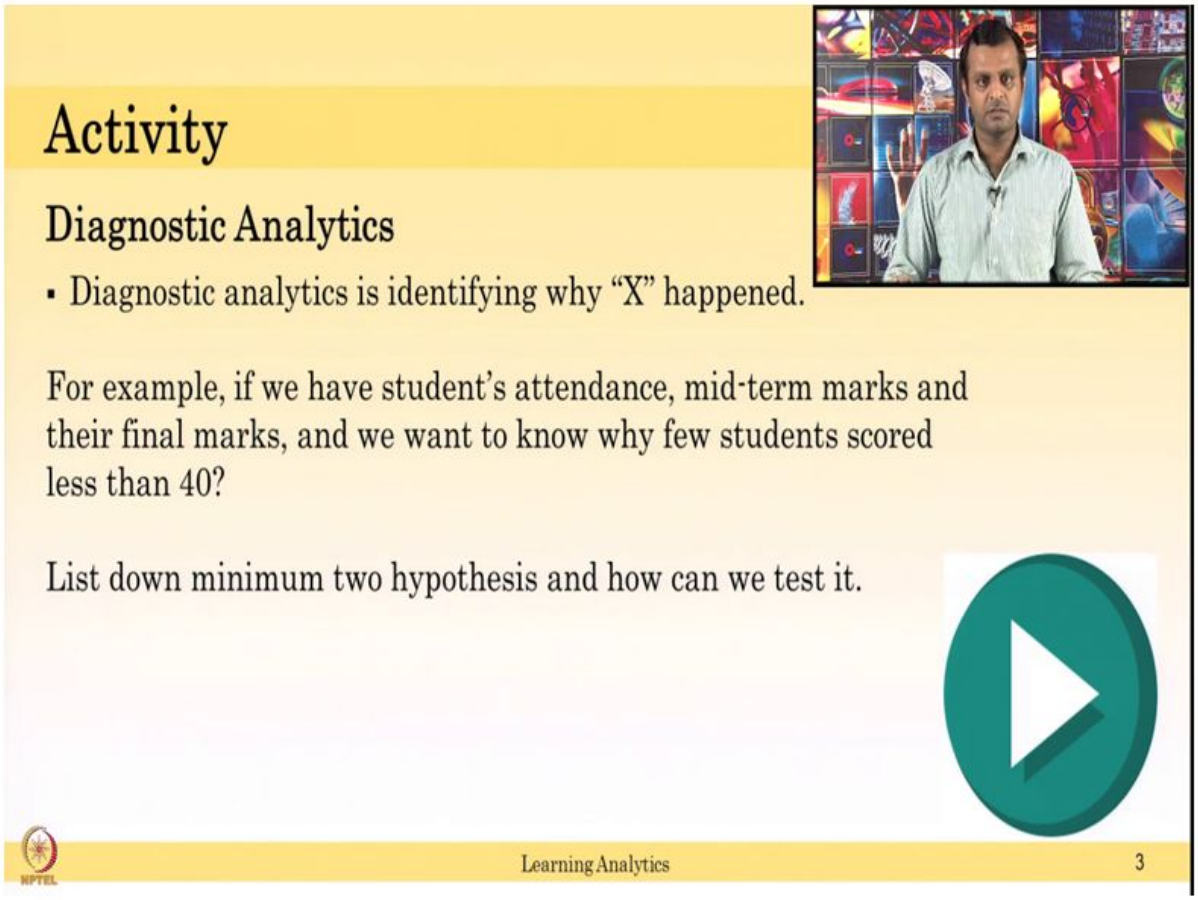


Hello everyone, welcome back to learning analytics tools course. In this week, we will discuss diagnostics analytics. If you remember this picture from week one, we saw that there are 4 levels of analytics, descriptive, diagnostic, predictive and prescriptive. In the last week, we saw

descriptive analytics, what data to collect, how to represent and that is very basics we saw because we do not want to teach a course on designing, better visualization or infographics.

But I hope you got the basics of what is descriptive analytics. In this week and coming week, we will talk about diagnostic analytics, then we will talk about predictive analytics. What is diagnostic analytics? So, it is to understand or to identify why this particular thing happened, why the students in the year 2018 are not able to score good or why the students in 2016 performed above average, why it happened? What was the factor impacting it? So, to identify that is called diagnostic analytics.

(Refer Slide Time: 1:22)



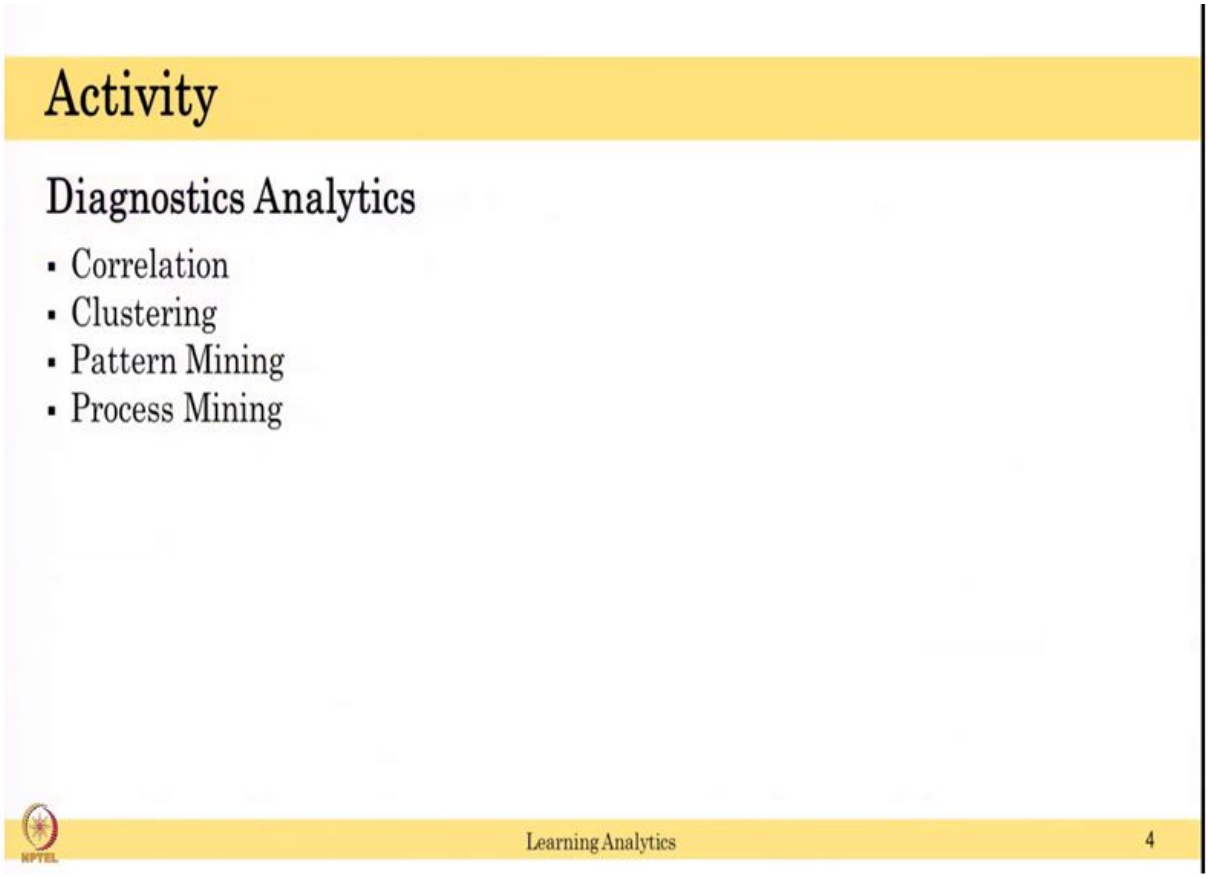
And so we saw that diagnostic analytics is identifying why a particular thing happened. For example, if we have students attendance and midterm marks and their final marks, let us consider this example for this activity. So, we will look at this example often to explain what is diagnostic

analytics. For example, if you have students' attendance, midterm marks, also the final semester marks, the final marks is the dependent variable, attendance and midterm are the independent, that is x1, x2, the final mark is y1 and we want to know why few students got less than 40.

Here, the focus is not who scored above 90, instead, the focus is on who is scoring less than 40. So, list down based on your experience why would that happen, so consider you have only these data, like midterm marks and attendance, can you create a hypothesis from these data and if you have a hypothesis, how would you like to test it? What are the techniques used to test whether your hypothesis is right or wrong?

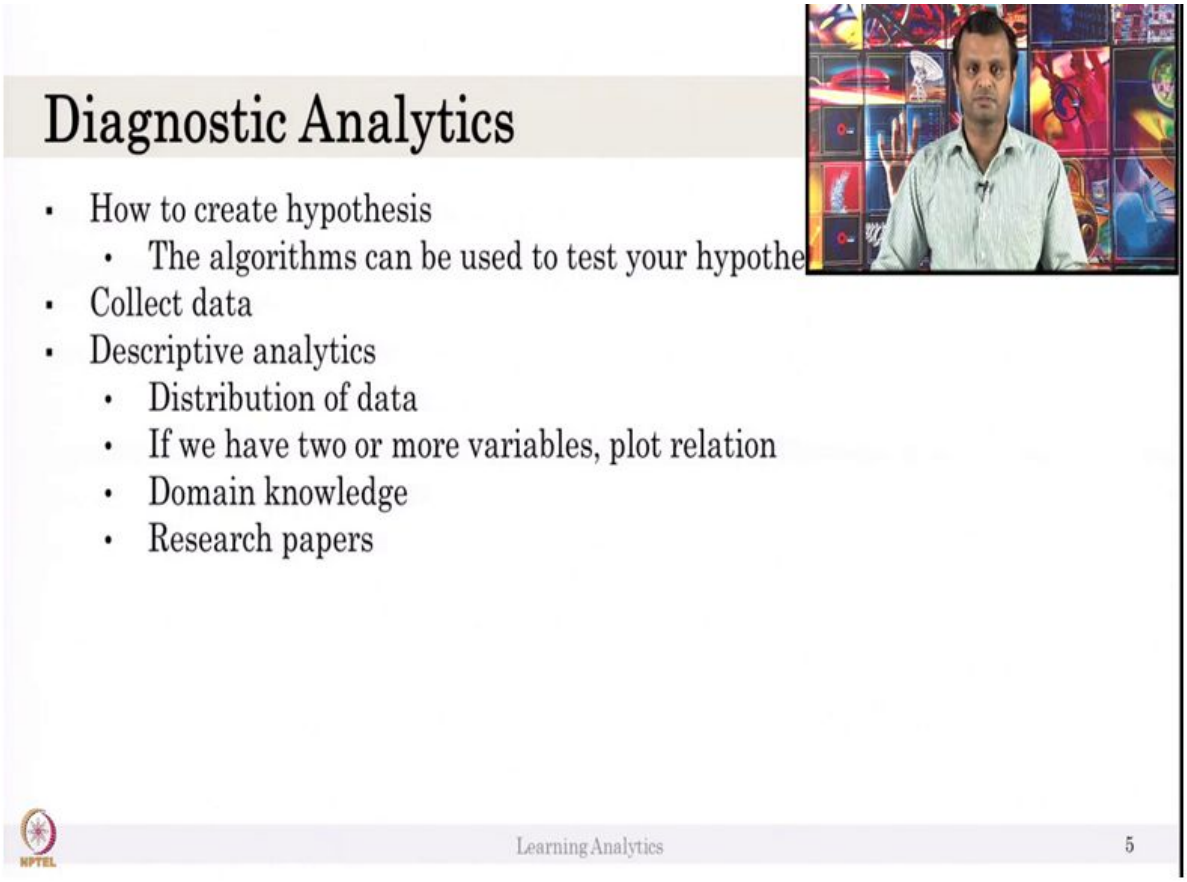
You consider that you have 1000 student's data or 60 students' data, like the number of data is up to your choice. Think that you have this data and you have x1, x2 and you want to identify the hypothesis, the relation between x1, x2 to y1 and come up with a hypothesis, also come up with a methodology to test it. Please resume the video after you get your answers.

(Refer Slide Time: 2:53)



So, there are a few techniques for diagnostic analytics. These are correlation, clustering, pattern mining or process mining. For example, this clustering or process mining can be also converted into credit analytics, but let us consider this for diagnostic analytics. So, the very basic step is a correlation. So, what is correlation and what is clustering? We will discuss the correlation in this week. In the coming weeks, we will talk about pattern mining, process mining, then clustering.

(Refer Slide Time: 3:27)



So, how to create a hypothesis, this is very important because you have data, like dependent data, also the independent data, how do you come up with the hypothesis. The algorithm can be used to test your hypothesis, but the algorithm will not create a hypothesis for you, write. So, the algorithm cannot create a research question for you. The algorithm can be used to test research questions or the questions you have.

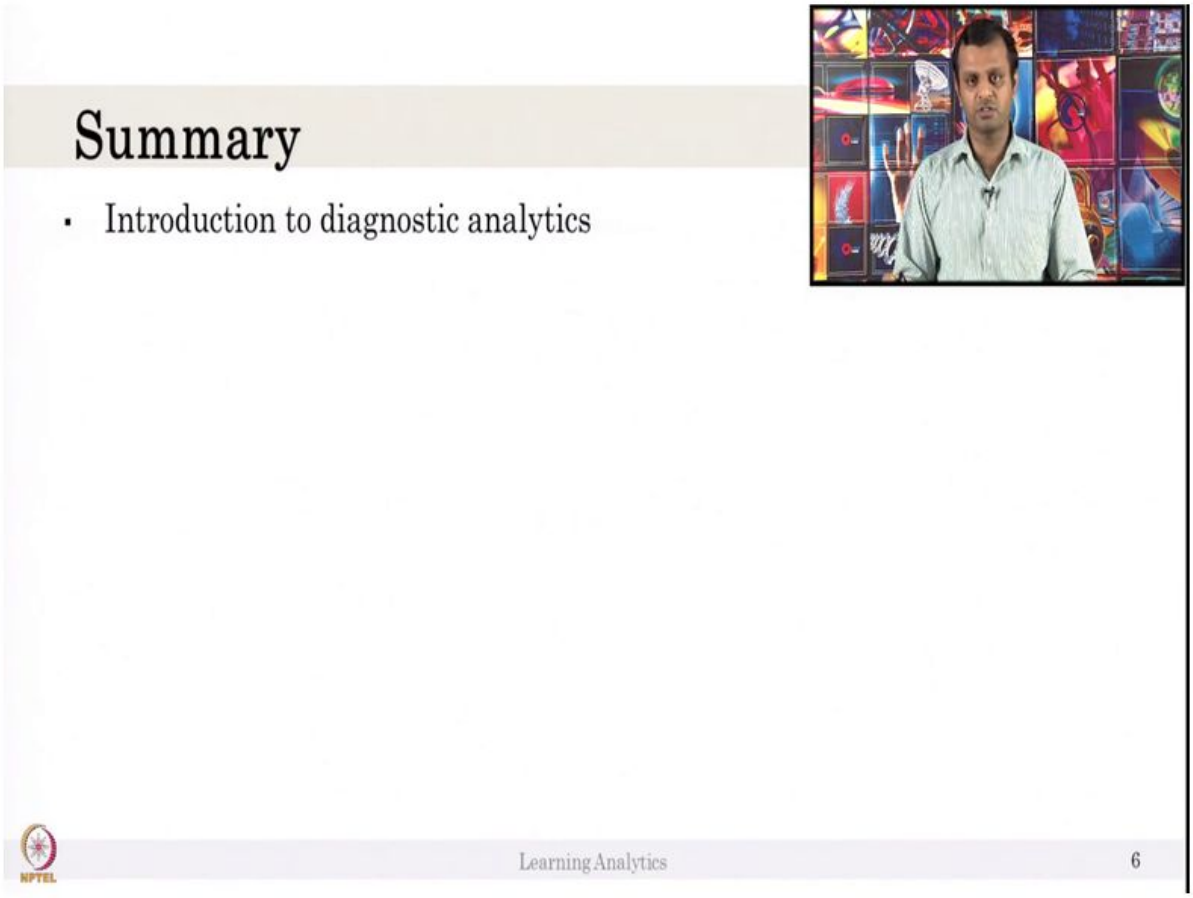
So, how to create it, let us start with the first step to collect data, like what data you can collect. After you collect data, do the descriptive analytics, see, visualize any connection, any relationship between this data, plot in a bar chart, distribution chart and in fact you go for a scattered lot, see the relationships.

Then, you have to use your domain knowledge to come up with your hypothesis. That is when you apply "learning analytics". So, your expertise in the domain helps you to identify the hypothesis. For example, you know that students who are not able to attend a class or were not able to submit assignments on time will not be able to get a good score in the final exam.

If you know that hypothesis and you have the data, if you want to test it, go ahead and test it. So, that kind of hypothesis, that kind of such question has to come from the domain expertise. If you are not able to identify such questions or develop a hypothesis from the domain knowledge, please read research papers.

The recent research papers in the related field and see what kind of questions they are coming up with and how they are creating the questions. You can start with testing the existing questions available on your data, then that kind of practice will make you think how to create a new research question. That is how you do research.

(Refer Slide Time: 5:26)



So, in this video, we just got an introduction to what is diagnostic analytics and in the coming video we will talk about correlation in detail, what are types of correlation, thank you.