Data Analysis and Decision Making - III

Professor. Raghu Nandan Sen Gupta
Department of Industrial and Management Engineering,
Indian Institute of Technology Kanpur.

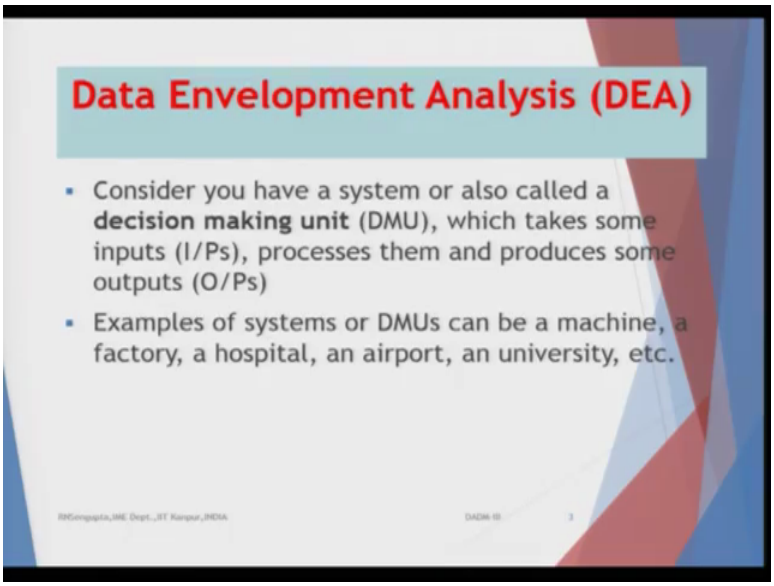
Lecture 36

Welcome back my dear friends a very good morning, good afternoon, good evening to all of you wherever you are in this part of this globe and as you know this is the DADM free course which is data analysis and decision making 3 another Nptel MOOC series and this total course duration is basically contact hours is 30 which when broken into number of lectures is 60 in number because each lecture is for half an hour and that is spread over 12 weeks and after each week you have each week consisting of 5 lectures which is of half an hour each.

You have one assignment and total assignments would be 12 in number as you can soon come to know in the next slide, we have completed 7 weeks which means 5 assignments are over and we are going to start the eighth one and by my good name is Raghunandan Sen Gupta from IME department at IIT Kanpur.

So this is the lecture number 36 which here is the eighth week and we will discuss something to more with linear programming for an application site and then go on to the nonlinear part. So this concept of data and enviorment analysis we have considered in DADM 2. I will go into the details, before if you remember we did discuss about the formulation on the idea that how you solve it using the operation research techniques would be considered now in a little bit more details. But the mythology of solving operation research or simple optimization problem using the linear programming you are already aware.

(Refer Slide Time: 2:01)



So consider you have a system or consider you have some decision making units, so why DA is utilized data envelope and, that means you have some datas based on which you formulate an envelope in generally in a 2 dimensional space in a Cartesian coordinate and what are the x axis and what are the y axis I want to come to that within few minutes.

To see for example, you want to analyze that how a hospital is doing or how a set of hospitals are doing or how company set of, factories of a company is doing or say for example how the government is functioning, different state governments in India are functioning or how different type of IIT's or IIM's taken collectively or as, as different entities are doing. And when we consider that how they are doing you are trying to basically find out that considering the overall set of inputs and overall set of outputs that how well we have been able to utilize this inputs and outputs to arrive at some rational decision. Rational decision means about the ranking.

Now these considered that you have the IIT's and if you consider only the, the first earlier IIT's and considering there are Khadakpur, it is Kanpur, Madras, Delhi, Bombay then you consider Roorkie and also you consider Guwahati. We want to basically find out that considering the inputs, what are the inputs here the question would be, would be there say for example a number of faculty members, the training on the faculty members how many papers they are publishing, what is the total number of grants they are getting, what is the total quantum in the money, monthly term, the MDP's they are doing, what type of consultees they are doing.

Then what type of teaching they are doing. Inputs can be seen for example the students, the ranking of the students, the quality of the students considering from which background they are coming. It can be the infrastructure, it can be the input how good the libraries, how the research labs are, how good the core labs are. It can be say for example how good the environment is, what type of recreation facilities there are, the staff could also be considered as an input, how efficient they are.

So, whenever you are considering the inputs there would be a set of inputs which are parametric in nature. You have basically some equational formats, some monitory values some units are there and you can basically find the total input in quantum terms. But there may be some inputs which you cannot find out in quantum terms is that if you are considering, say for example the type of students who are coming.

Leave aside their ranking in the JEE main or the JEE advanced. You will, you will consider that they are coming from the different state of the society, from different economically poor backgrounds. So obviously we will consider that as some subjective input which would definite be considered, in a, as one of the inputs when you are trying to basically collectively find out the input bundle in order to basically rank this different types of IIT's or the decision making units which we call the DMU's.

Now, on the other hand the outputs can be, say for example what is the total number of companies which you are visiting, the institutes, what is the median salary, how many students have been placed, what is the Indian, type of Indian companies which are recruiting people, how many female candidates have been placed. Obviously we can consider the female candidates as an input also considering that there is a thrust area in the education system to basically put more weightages that number of female students who graduate in that technical education should be much higher.

Then the output can be what is the research quantum of the students who were publishing. The quality of M-tech thesis, whether M-tech thesis are being published, whether the students are going for higher studies, whether the students are going for, say for example PhD students going for post doctor position, what type of employment these PhD students are getting into different IIT's, NIT's and technical CFTI's.

So, here also the number of outputs can either be quantified in monetary terms, in some units, in some mathematical terms and there would be some outputs which cannot be done unlike wise so obviously you will have some sets of outputs which are subjective and some sets of outputs which are objective. So, you would basically like to combine both the inputs subjective objective to get some output which is subjective, objective and then basically combine the inputs and combine the outputs.

Try to find a ratio, try to find an average and then rank the inputs and the outputs combined together such that we get a ranking system. Now as I said that this type of data environment analysis can also be utilized for hospitals. so the hospitals would be, say for example what is the number of beds which is the hospitals has, type of, different type of operation theatres which are there, what type quality of services they are, what is the ICU numbers of patients who can be admitted, what are the type of nurses, whether they are well trained, what are the types of doctors, number of doctors, what are their degrees.

Then inputs can be how many people are being admitted on daily on an average basis and the outputs can be, like say for example it can be number of successful operations which are there, number of successful complicated medical cases which the hospital has been able to diagnose, or the average number of stay or general patient has to stay in the hospital because lower number of days a patient has to stay and he or she requiprates and leaves the hospital it is a positive mark for the nursing home.

Then it will also mean that what are the number of different number of patients from different state of the society, the hospital is able to cater too. Whether the it may be also be possible that what is the total amount on an average which required for say for an example in appendix operation or any heart surgery, if they are very low obviously it would mean that many of the people who are being admitted in general form from the society would be able to meet the cost in order to be medically treated.

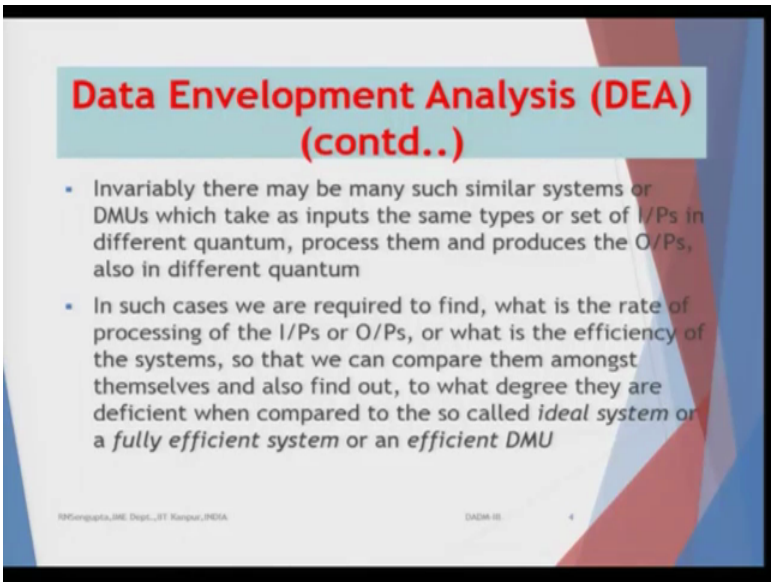
And, similarly we can give examples for the government so it can be say for example number of outputs would be in a type of road feature the electricity has been supplied, the drinking water, education inputs, may it be taxes, may be efficiency of the administrators, the output can be severing how fast the garbage cleaning system is there, so and henceforth.

So in the first example which I gave for the school, second for the hospitals and third for the government institute are the state governments or the municipality or the corporation. We have, whatever examples are inputs and outputs we have discussed, there would be some sets of inputs which would be subjective, some set of inputs will be objective, similar would be the case for the outputs which would be subjective or objective. So we are basically trying to analyze some long parametric methodology of trying to rank different above DMU's and here we will see how we can utilize the concept of linear programming and optimization to solve such simple problems. So with this background let us come to the slide.

So it says consider that you have a system or the so called decision making units or DMU's which takes some inputs, bundle of inputs processes them, whatever the processing things are for a factory it would be doing some manufacturing process. It can be grinding, it can be shaping, it can be like trying to basically do the different type of smithy work and all these things. In the school it can be training with the student undergoes the teaching of the subjects which the student undergoes.

And for the hospital it would be basically be the different type of medication, the patient undergoes so and henceforth. So once you process them you produce, you produce some outputs and the outputs I have already discussed. So, examples of the system or the DMU's can be the machine, a factory, a hospital as I said an airport, it can be a university as I have said the IIT's starting from Kanpur, Kharagpur, Kanpur, Madras, Mumbai, Delhi, Roorkie and Guwahati.

(Refer Slide Time: 10:39)



Invariably there may be many such similar systems or DMU's examples which already I have given. It can be the banks also like if you are trying to basically consider the Indian public sector banks so need not the banks, the main motive need not be profit it may and main focus would be to cater to the mass population in the country where they want to spread the concept that people can deposit their money in the banks put and go into fix deposit go into a recurring deposit and utilize the loans like car loans, personal loans, house loans, education loans in order to basically understand that how the, utilize the banking system to the maximum possible extent and then when you are trying to basically consider for an example SBI or PNB or Allahabad bank.

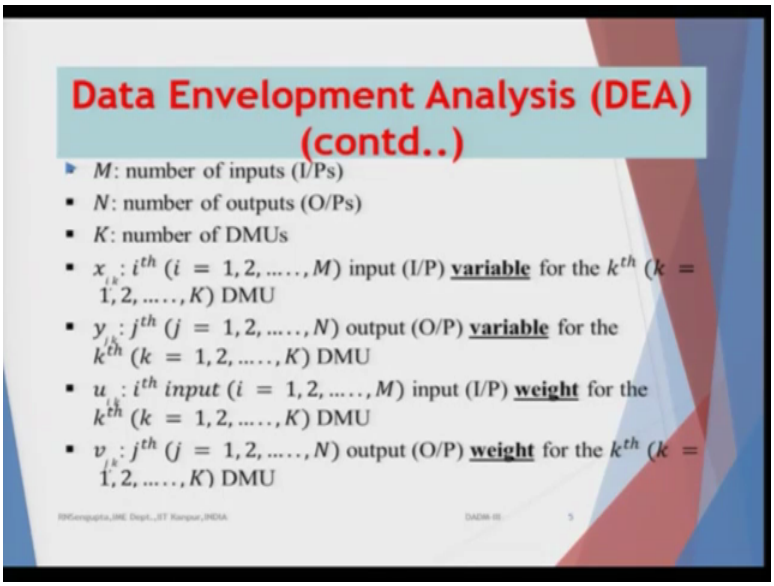
So there the profit may not be the main reason based on which you try to rank the banks. I am talking about the public banks, public sector banks. So invariably there may be many such similar systems or DMU's which take as input the same types or the set of inputs in different quantum, in different numbers of values, process them and produces the outputs also in different quantum's. In such cases we are required to find out what is the rate of processing out these inputs or in order to produce the outputs or what is the efficiency of the system.

Efficiency means in general terms you will use the gen very simple definition of efficiency also. So that we can compare them, compare them means the DMU's amongst themselves and also find out to what degree they are deficient or, or there is inefficiency in the DMU's and when we are trying to compare these DMU's with respect to the so called efficient one's, so obviously

there would be a lack of efficiency, fall in the efficiency, we will try to basically compare the inefficient DMU's with respect to the efficient DMU's.

So that means let me read it, so its main study would be to what degree they, they means the DMU's which are deficient or inefficient, so they are deficient when compared to the so called ideal system or the fully efficient system or an efficient DMU's and how we compare them. So we will basically have a nomenclature.

(Refer Slide Time: 12:56)



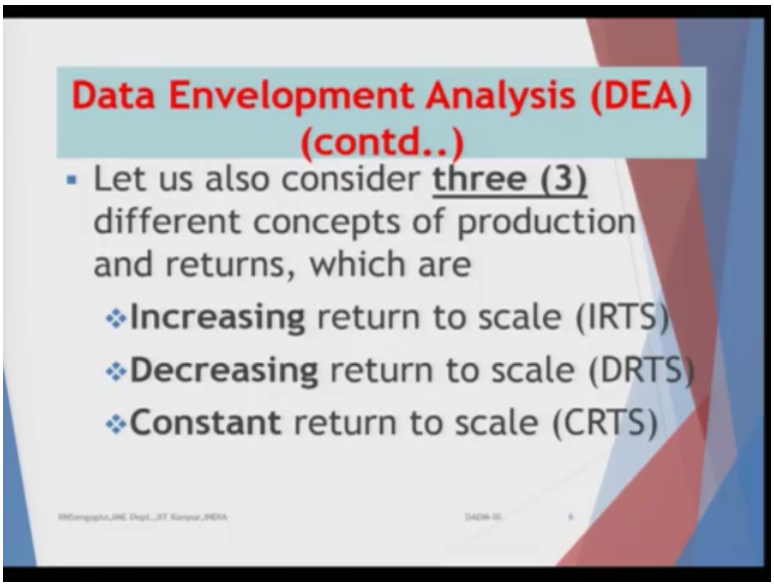
So what is the nomenclature we will consider the number of inputs. Distinct inputs are M in number, M as in Mangalore or in Mumbai the N would be the number of outputs which N as in nose or in Nagpur. K would be the number of DMU's and x would basically be given by the concept of x suffix ik where i would be the ith DMU where i is equal to 1 to M. where it, it will 1 to M, not the DMU to be the input concept. So i 1 to M would be the input variable for the k th DMU where k is equal to 1 to k and this i, m i j and k are all small while y suffix j k would be the corresponding to the j th, j is equal to 1 to N output or variable for the k th DMU where k is equal to 1 to k.

Correspondingly we will have weights, so the weights to be given as u i k, which would be the corresponding input weights corresponding to the x i k decision variables and similarly for the output would be, you will also have v j k, suffix j k which is again being 1 to n and k is equal to

1 to capital k and these would be the output weights corresponding to the y j k output decision variables which we have.

Before we go into the concept of trying to basically analyze the problem in very simplistic operational research framework and optimization problem so we will consider 3 different types of concept of production returns. So what about production returns, what about economic environment you will have?

(Refer Slide Time: 14:47)

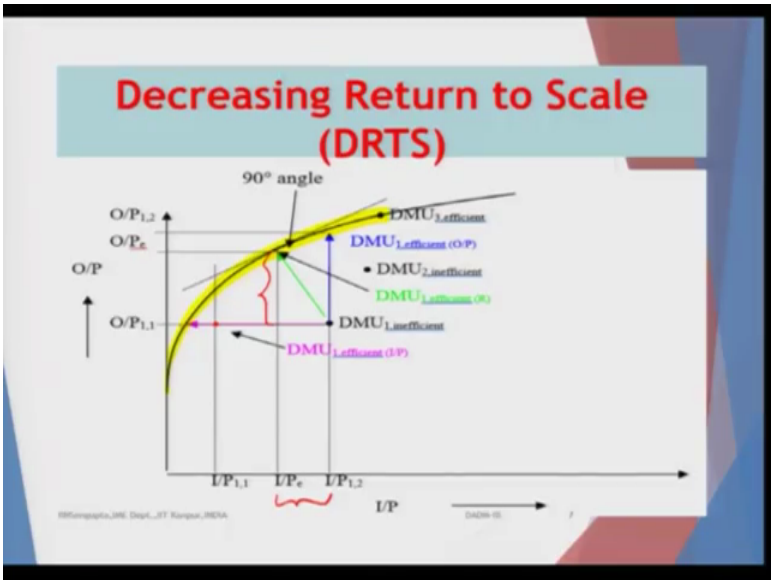


So one will first we will basically mention as the increasing return to scale that means the more you produce per unit, more is the per unit profit you are going to get. So you will basically try to increase your production to higher and higher levels because the per unit profit will keep increasing. So keep increasing the word which I am going to, which I just used will make sense when we come to the simple mathematical concept which we will just discuss.

The second concept is the decreasing return to the scale or the DRTS where more you produce, per unit output will start decreasing and this would be in the areas when economy is not doing well. Any extra production would only increase your fixed cause. The returns would start falling as you add more and more units to the sales. The last part would be the constant return to the scales which is CRTS and here you will consider that any additional input which you are going to make, the quantum of additional profit or revenues, or revenues minus cost whatever you think it is per unit return you will get would be on a constant scale.

The first is increasing per unit time addition, second is decreasing per unit addition and third one is constant per unit addition. So you will try to basically analyze RDA concept on under this framework. But before going into the increasing of the DA framework or D optimization problem in the simplistic sense, I will basically go through simplest, very simple diagrams for the increasing return to scale which is IRTS, decreasing return to scale which is DRTS and constant return to scale which is CRTS. So the diagrams would basically be of the same type but only the orientations will change and the graph, overall shape will change and that would basically give you that whether it's increasing, decreasing or constant. So in this case just concentrate on the x axis.

(Refer Slide Time: 16:47)



This has been run in the DADM 2 but as it is a new course hence I will go in the same level of depth as I have done for DADM2. So in considering back, coming back to the slides along the x axis, you consider the input or the input bundles so IP, along the Y axis you consider the output which are the output bundles and the graph which you see which I have mark now and use the highlighter yellow one.

This is the graph which you see corresponding the case where you have the decreasing return to scale. Why decreasing return to scale. So let us concentrate concentrate on IP suffix 1 1 IP suffix 1 2. So or IP suffix 1-1 if I go, I am not going to draw it here because that will become too clattered so if I go up, vertical up, so IP 1-1 basically means this so called frontier at this point

and if I go vertically, horizontally left I basically come to an, the this so I should basically be mentioning this point later on.

So considering the input comes here which is the point and when I go to the left it basically reaches the output level 1-1, so generally the input output bundle is such that the DMU which I am marking here is below the, the efficient frontier so we will basically consider it as inefficient. Why we are going to consider inefficient because if we go vertically up we are able to reach an high level of output based on the same level of input which we have or else we consider that we can maintain the same level of output if we move on to a lower level of input, if we move on to the horizontal on the left, left means from as you see from the graph.

Now consider the, the IP 1-2, where you have again the inefficient one, and the inefficient one basically goes up to this point. So IP 1-1 I just do it in order to make things clear to you. So consider the DMU 1 which is inefficient. If you go vertically up the blue line you reach the, the efficient frontier. So hence at the same level of input you are trying to increase you output.

If you go horizontally left you will reach the pin dotted lines, if you consider you will reach this point where the electronic boundary is highlighting, that means at the same level of output you have been able to decrease the input to the maximum possible level. Now consider another DMU, which is basically IPE, E means efficient. If you go vertically up, it will obviously touch the efficient frontier at this point which is where the green arrow is being basically pointing. This is just a coincidence that, this point should have been in this direction so the efficient one so if it goes up, so if you go vertically left again the output level is efficient.

Now, if you consider the DMU, so, why this graph is called decreasing return to scale, so if you consider per unit increase in the input bundle the height of increase or the output is increasing, but it is increasing at a decreasing rate. So we will consider the d by dx is positive but the second derivative obviously would not be positive, because in that sense, that per unit additional input which you are adding, will increase your output, no doubt but the increase would be happening, happening at a decreasing rate.

So here the graph is basically following. So the d by dx would definitely be increasing but increasing at a decreasing rate. Now if I consider, the inefficient DMU, so before going to the

next graph, so if I consider the inefficient one which is the back one, so there are 3 ways it can reach the efficient frontier. There are many ways but I will consider only 3.

One is basically going vertically up, one is basically going horizontally left and one is going in the tangential direction towards the graph. So what does the vertically up mean. The vertically up mean. That means that I am trying to maintain the same input but I am trying to increase the output. If I go horizontally left it means that I am trying to basically maintain the same output but I am trying to decease the input and If I go in the so called radial direction or tangential direction, this DMU, this inefficient one will try to basically reach this graph where the electronic pointer is highlighting says that it is tangent to the graph.

In the sense that there would be some proportional decrease in the input, proportional decrease in the increase in the output also because the output initially was this so the output would impleased by this amount and the input would basically decrease by this amount. So that we will consider in a, in a scenario, these 3 things will understand in a scenario where in first case we are only concentrating on the output without touching the input. In the second scenario we are only concentrating on the input, without touching the output and in the third scenario which is the green one tangential movement.

Can tangential movement because that is the shortest distance between the inefficient one and the, the graph which you have. So it will be moving in such a way that the proportional in decrease and increase in the input and output would be such that you would basically be maintaining the same, high level of efficiency and hence you are on the official frontier. Now when you are considering this 3, movements, there is the blue one, the pink one and the green one they would basically be utilized in 3 different environments.

So If I am considering only the blue movement vertically up, you will think that here the input cannot be touched, so you have to only concentrate on the output. In the pink one where you are moving horizontally left here you will consider that the output cannot be touched, only you can increase the input and then in the green case you can basically make a so called compromise where you will decrease the input also as well as increase the output.

Now where these problems would be applicable is then. In the case say for example you are, you have taken some labor and the labor has all, I will give you an example of the labor of the ones

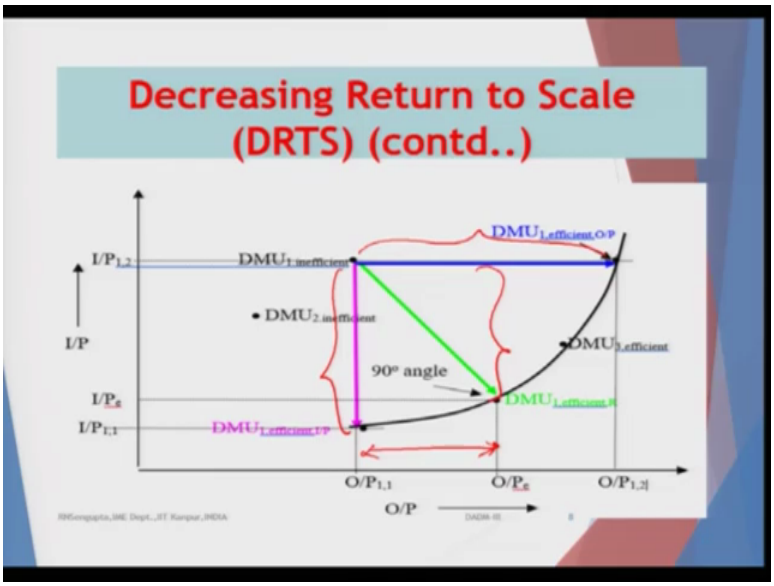
also and the machine ones also considering the laborers are all permanent staff so if they are permanent staff you cannot fire them. But what you can do is you can retrain them, train them such that at the same level of input you are basically trying to, input means the number of laborers, the amount of cost which you are going to give to the laborers, the salaries so and so for without putting any extra overtime, you will try to basically get the output in and at a much higher scale.

Or say for example you have a machine and these machines are, some of them are not working and you will basically repair them such that you will try to get the maximum benefit from the same set of inputs, because the inputs cannot be thrown away. That means you cannot fire a permanent worker.

That is where you are going to consider the vertically upward line. In the other way when you consider the horizontal left direction you will only consider the output where you can decrease the input so decrease in the input can be say for example, you have some permanent staff and some temporary stump. So this temporary staff would not be hired anymore they would be fired depending on the labor law which is there in the state or the place where these DMU's are operating considering that I am only taking the example of, of a factory or DMU which produces some outputs considering some set of inputs.

So you will consider either doing away with some old machines or trying to basically fire some temporary workers as that you maintain the same level of output by trying to basically get or utilize the inputs at a much higher level scale. While if you considering green lines it would be combination such that you basically decrease the input like fire some workers and trying to increase the output by giving training to the extra people who are there.

(Refer Slide Time: 25:53)



Now consider the graph where you have, where you basically switch the x axis and y axis, so here along the x axis you have the output, output bundles and along the y axis you have the input or the input bundles, so it would be again denoted by IP 1-1 then IP 1-2. This suffix 1-1 and 1-2 are the suffix as correspondingly you have the output which is OP 1-1 and the OP 1-2. Now remember one thing, here the DMU's which are inefficient they would basically be decreasing the input if you go in there, in the vertically downward direction.

There the output would increase in the earlier graph it was increasing if you are going vertically up, here the input is decreasing if you going vertically down. And in the other case in the earlier graph if you went horizontally to the left you basically decreased your input and if you go horizontally on to the right you increase your input. In the same way if you consider this graph, If you go horizontally left that means you will be decreasing your output and if you go horizontally right you will be increasing your output.

So if you consider the DMU which is inefficient, again which is this graph the blue, red, the black dot you can go vertically down, the pink one and reach the efficient frontier. So efficient frontier you are reaching in such a way that you are maintaining the same level of output but trying to decrease your input, because the decrease is happening here. If you go horizontally right you are maintaining the same input and going and trying to basically increase your output, that means you are trying to basically reach the efficient frontier by moving horizontally right

and if you move diagonally that means you go in the direction where you reach the efficient frontier in the shortest possible time the 90 degrees, obviously you will basically be come, combining the decrease of the input and increase in the output in some way.

So the output would be increasing by this amount, I am using the same color red, please do not mind and the input would basically decreasing by this amount. So you are decreasing the input and increasing the output. So this we will consider the decreasing return to scale where any additional, addition of the input would basically decrease the output by per unit in the, in some quantum, as I mention the d by dx is positive but increasing at a lower rate.

With this I will end this first lecture of the eighth week and continue discussing more about the constant return to scale and, and increasing return to scale and come to the formulation of the DEA say, explain how it is done as a ratio and then basically formulate, reformulate that as a linear programming and then we will follow the same procedure how you solve the problem using the linear programming problem. The tableau format so and hence forth. Thank you very much and have a nice day.