Course material

Course:

ENG606 / PHYS 442

Video:

DOE_lesson2_part2_Weighting3Objects

Concepts (extracted from automatically generated subtitles):

Design of experiment. Main factors. Variance of output. Mathematical model of your situation. Offsets of your measurement. First things. Matrix of the model. Different elements. Mind map. Lot of time. Most important possibilities. Lot of factors. Linear model. Different types of designs. Second degree.

to video sequence search
(within ENG606 / PHYS 442.)

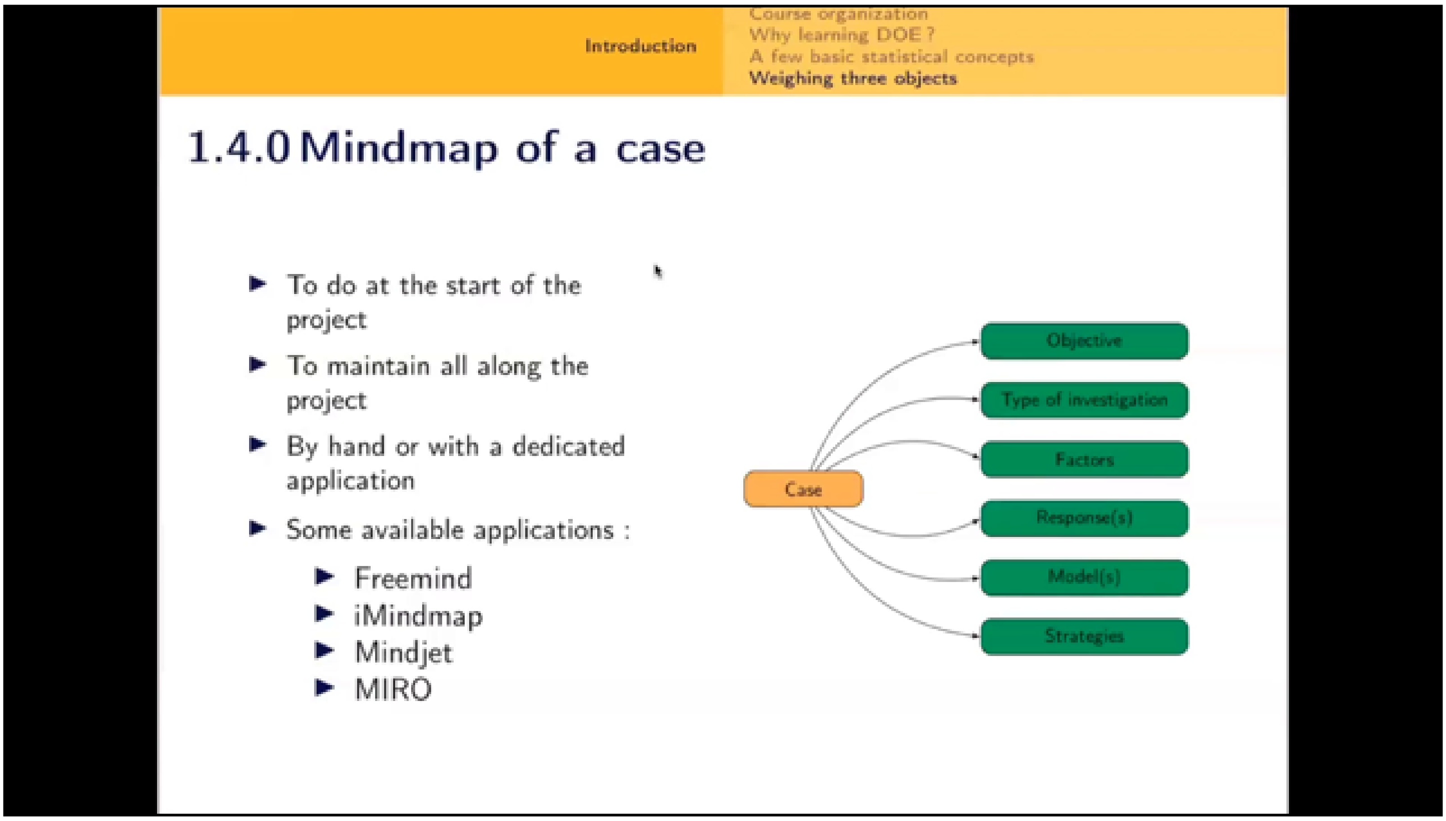
to video

...

notes

summary

0m 0s

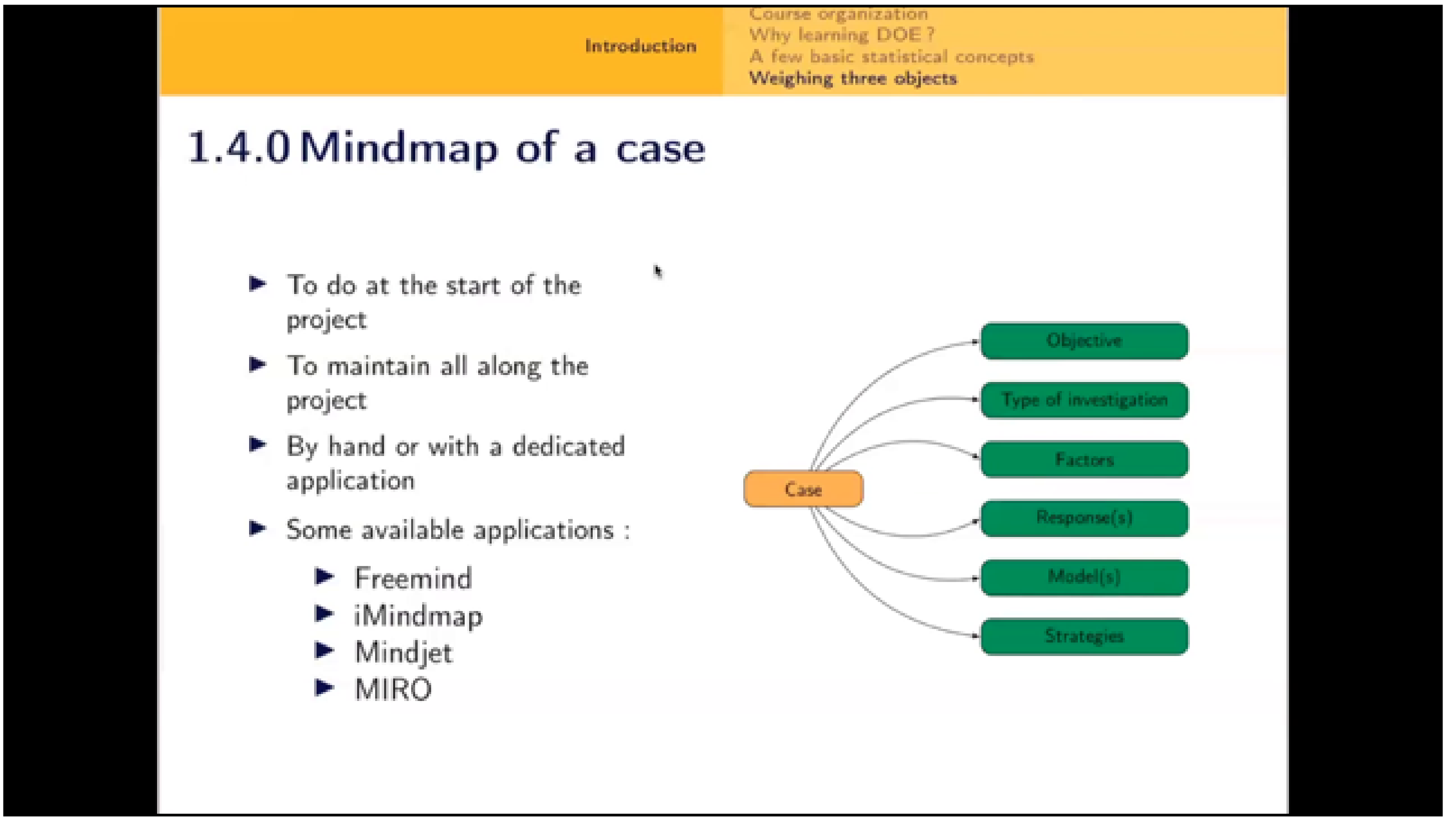


These subtitles have been generated automatically moment when you are performing design of experiment and more generally statistics, sometimes you are the engineer with the problem of technology, with the problem with the different elements that you have. And another moment, you should be a statistician, you are not interested by the sensing, you are not interested by the width of the probe, the temperature of the et cetera, you are just interested by what is the flow of information that you have to treat. So this mind map, you can make it by hand on a piece of paper, you have some tools like three mind in mind map, man jet, mirror, anything you want, but it's really important and I really encourage you to do it. As you want, if you come with your draw with your computer, in fact, those are the elements that you need to have clear in your mind when you would like to start making a design of experiment. So the first things you have to be actually clear is the objective. Why you are doing this? Because there are different types of designs, strategies, and you don't know exactly what you want, you will never really get where you would like to go. So they are very typical two objectives, you can decide that more precisely, but sometimes you would like to know what are the main factors, you have a lot of factors and you want to make a selection between what matters and what doesn't matter. This spirit is a parental principle that eventually a minimum of factors will have a maximum of consequences. And as a good engineer, this is the art, you look for the few factors that let you manage your your process. The superglant possible objective that you want to optimize, you want a minimum of something, you want the maximum of something, so to be

notes

summary

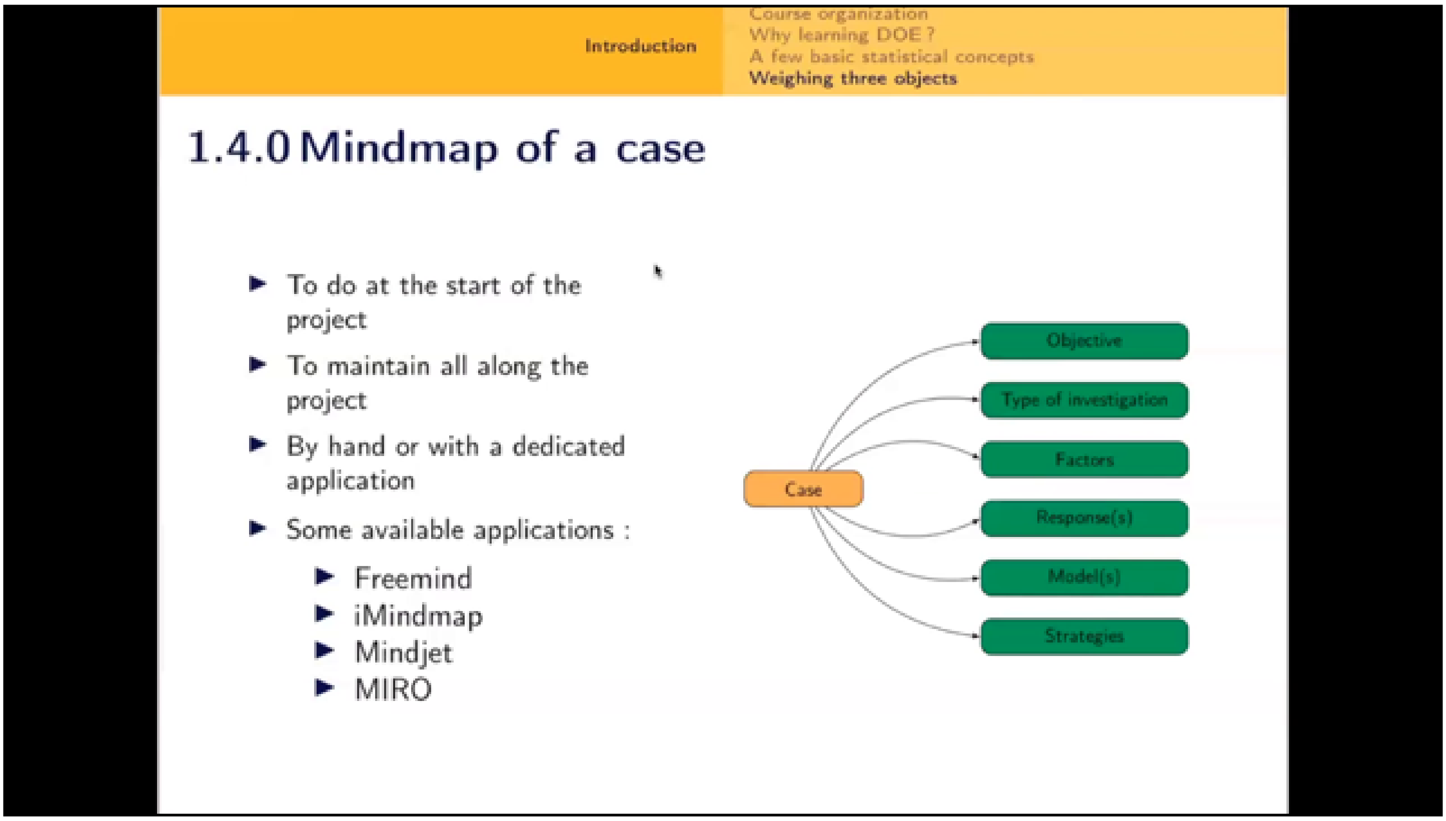
0m 1s



honest, it's making the food like to be a mathematical model of your situation and try to find in a surprise what where are what you look for. It could be maximum minimum, it could be also some past because you would like stability, sometimes not exactly the value of your output that is important, but it's the stability of your output, the variance of output. So they are very two strategies. So it's very important to have clear the objective. After I add this recent, is a type of investigation, usually in engineering, you're making experiments. This is the type of analysis that people in medicine, in biology, sometimes they make inquiry, sometimes they look for the official data and select the data that are important for them, they make laundry to denial and then these are different types. This is not experiment. We call it experiment when you have access to your factors and you can change that. It's what we usually made in a laboratory. So here typically type of information, let's say it's experiment or laboratory, you really have access to your parameter and this gives you freedom for planning what you want to do. That doesn't mean that you can do any value, sometimes you have limitations, some experiments could be very costly, some others would be explosive, I don't think that you can do all possible experiments, but I said that you can manipulate your variable. If you are making experiments with meteorology, you cannot decide which wins, you want to do something, but if you are making an analysis in situ of your meteorology, it's not experiment situation. After it's important to have the list of factors, the problems are more or less complicated from your factor. What they are, how many you have. So the first thing that you have, what are the range of your factor and after is if they

notes

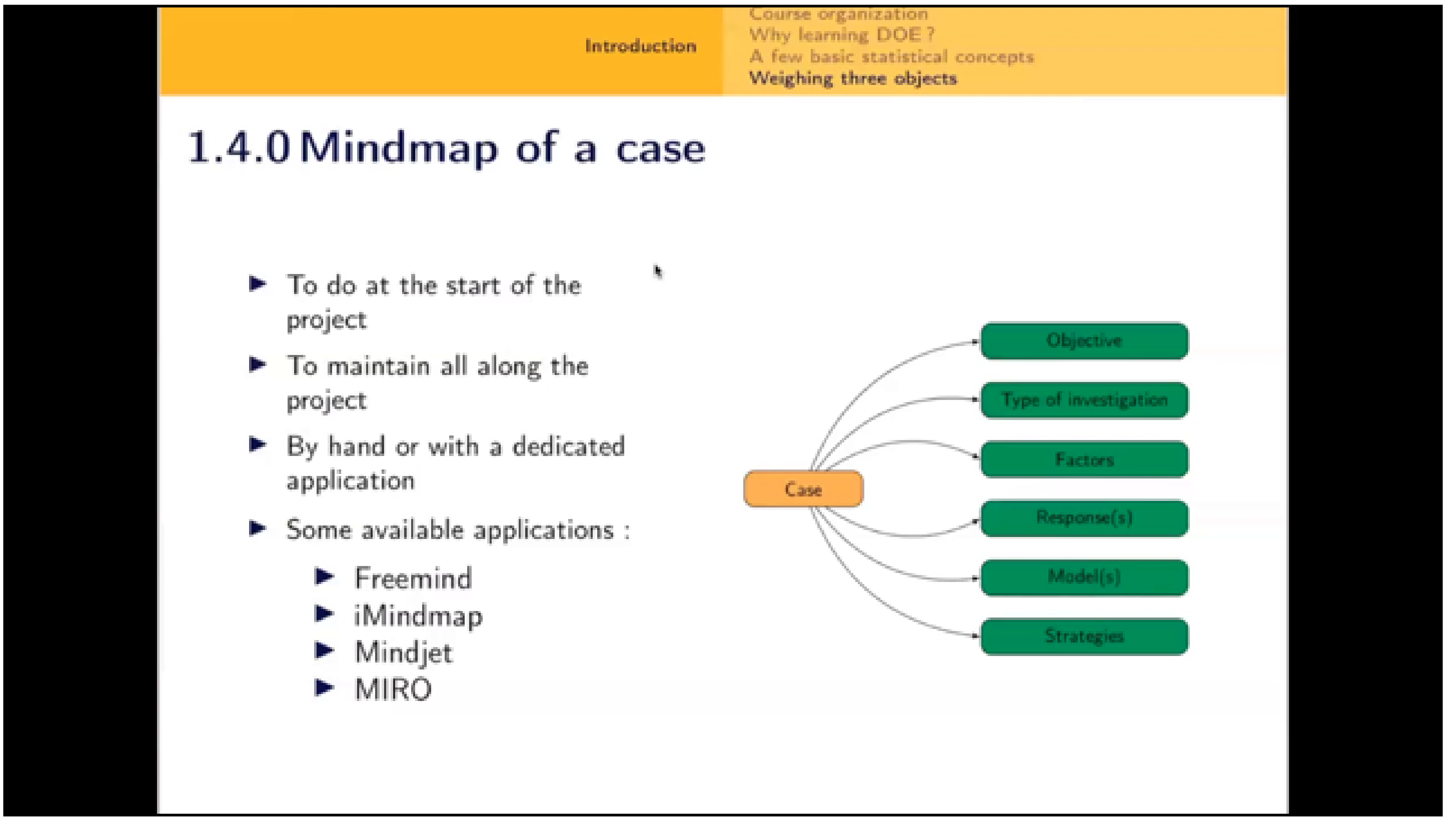
summary



are independent. So usually here I like to see the list of the factors, understand if they are qualitative or quantitative, do I have to choose between the project A, the project B, the project C or I have to choose between the sickness of 1 millimeter, 3 millimeter, 4 millimeter, we have two very typical types of factors. And if they are independent, because if they are independent, the space in which we will play will be an hyper cube. If they are not independent, we will play in space that are different. Most of the solution are geometrical solutions of the space in which we are playing is very, very important. After I would like to know something about the response or the responses. Initially we think one response, some factors in one response, we could have a situation in which we are checking for the price of something, the efficiency of something, the eco toxicity of something. So we could have different outputs. Sometimes they are so wide and show you some response. We can make a mixture of the different response for having a normal response taking care about the cost, the cost of the toxicity, about the easiness to produce something. So the response. What they are and after we have to think to the model. So what are you looking for? When you are looking for, what is important in a process, if you work with linear model with interaction, usually it is sufficient. If you are looking for the maximum or minimum, you absolutely need to have a second degree in your model because you need a heel and if you need a heel, you need a second degree. Eventually your model will not be a perfect second degree function. But what you want, you won't understand why you have to spend a lot of money. Eventually in the model, you have a

notes

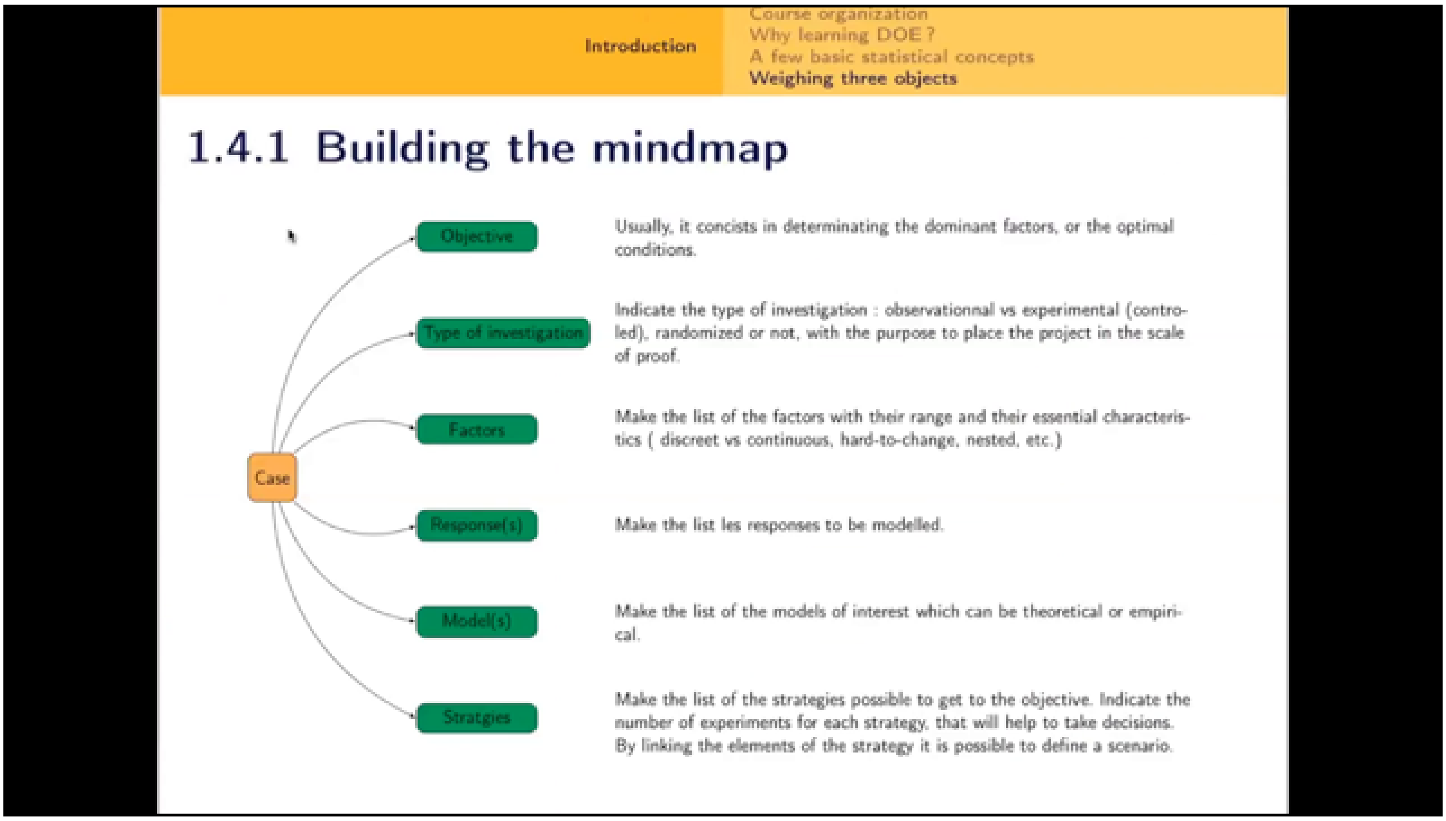
summary



model. You are working with a thermal model, which is known. You are working with split. Perhaps the family equation is really your model, you want to work with that. Or perhaps it's too complicated and you would like to work with a simple model. So it could be empirical, it could be theoretical, it could be something in between, but you have to precise in fact to what you would like to compare your data. And after only we start talking about strategy. And so the idea is having the list of all the most important possibilities and check, calculate which one is the most interesting. It could be because of the delay, of the cost, it could be for the accuracy. You sometimes have

notes

summary

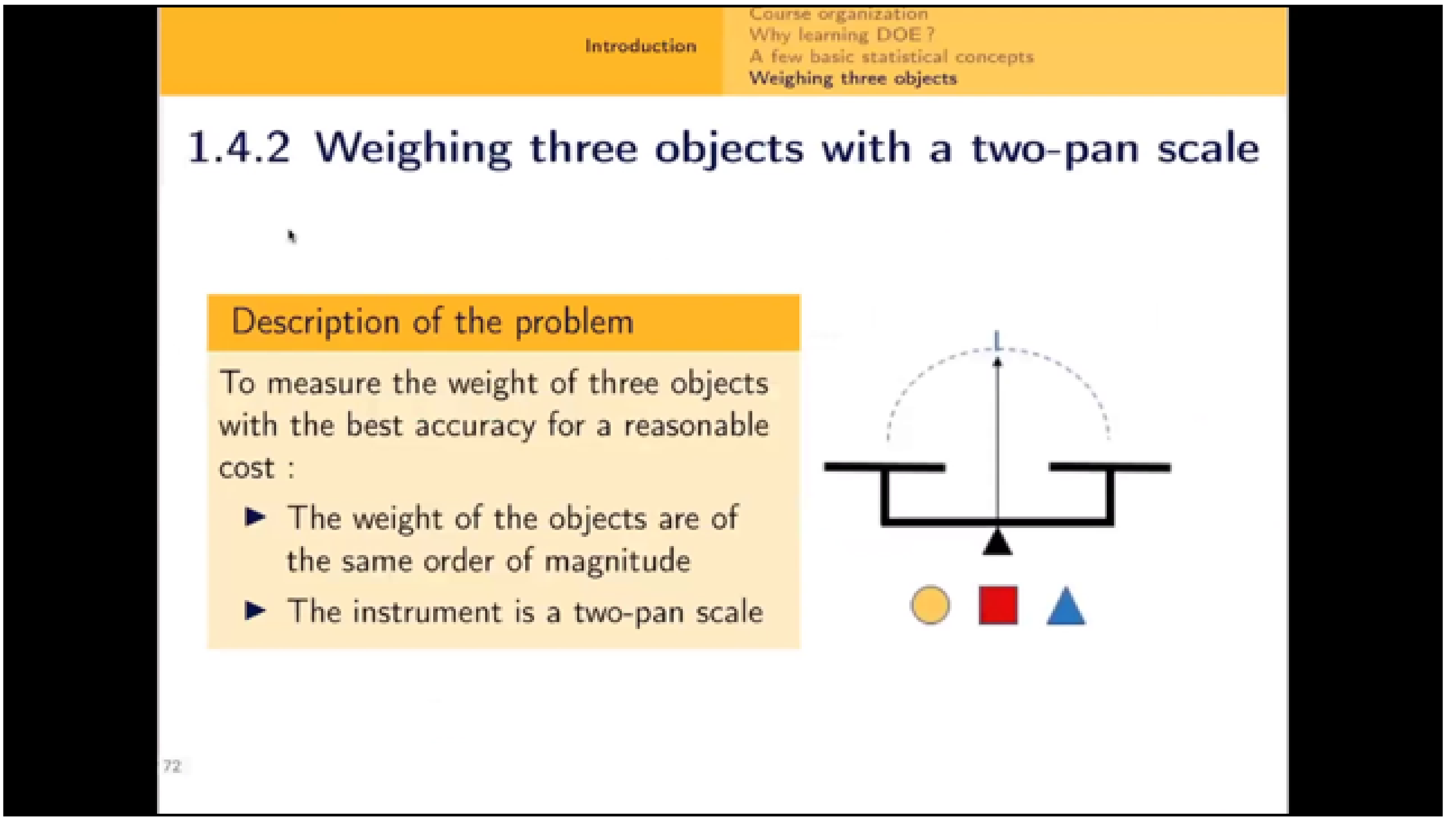


different models to say, okay, for the next paper, it's okay if I work with linear model subtraction. But at the end, I would like to go to a model with some surface curvature and I need more. So this is what is the strategy point. And after when you are making your experiment, you follow that. If you change points up there, wow, you're like, in new map, let's say, that after you just follow the different strategy, because here you can have a stupid gift, then you make the first experiment to get some results and you continue to get. I had 10 factors, I made the first experiment, I can take out seven of my 10 factors, I just have no three important factors that represent, I don't know, 80% of my process, or my process, and no, I go for second degree with my, with my own, because it could have been a lot of time to start from scratch going at 10 factors, two degree or certain way, so many experiments, it's not interesting. So it's why this map is very, is very important. And this map is for you, for following or your laboratory, not so, but it's also for discussing with your, with the statistician, the statistician out here, with the different partners, the stakeholders of your experiment, it's well with

notes

summary

8m 25s

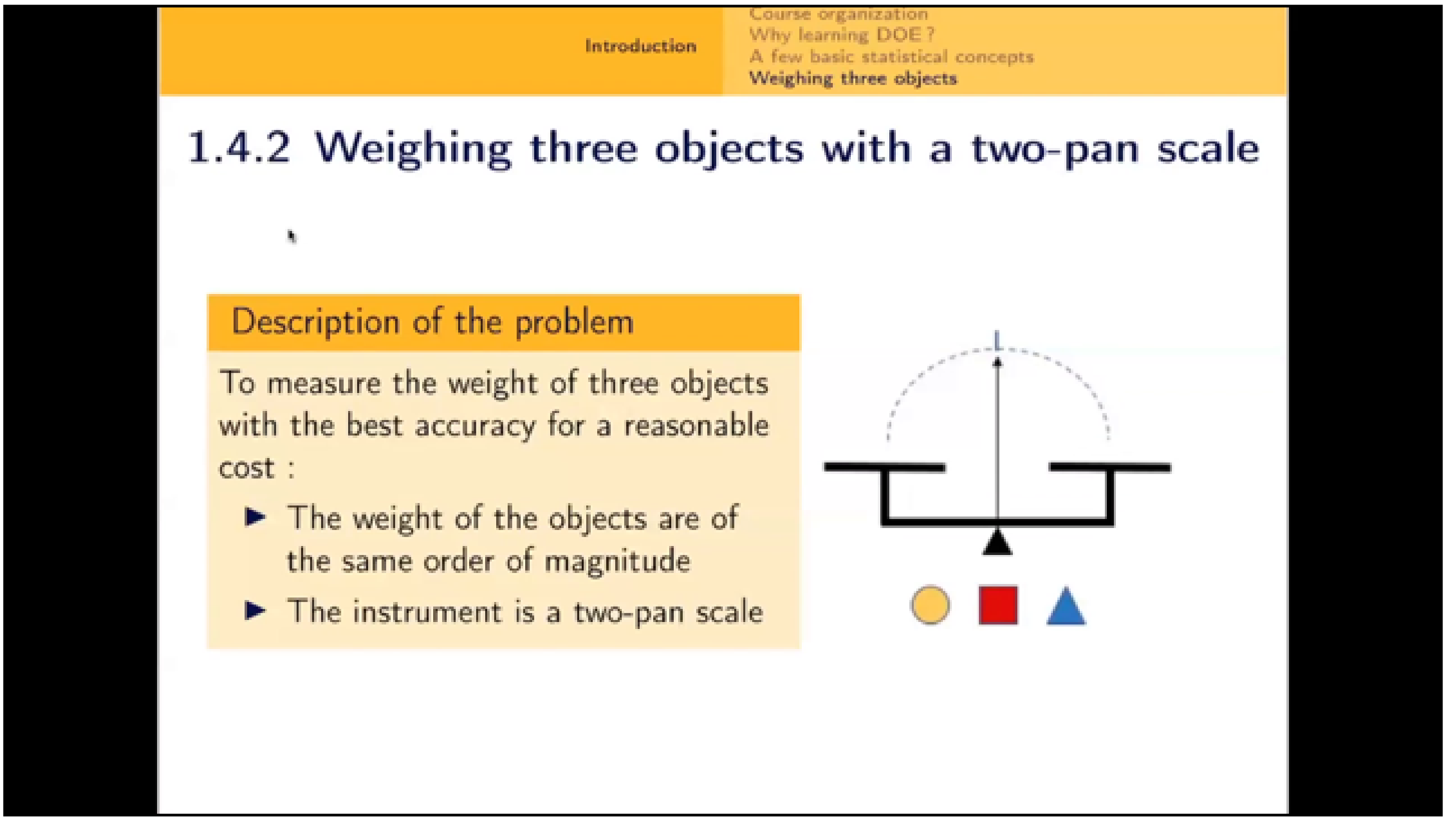


the backbone of your experiment. So here you have a slide explaining what I explained you

notes

summary

9m 57s

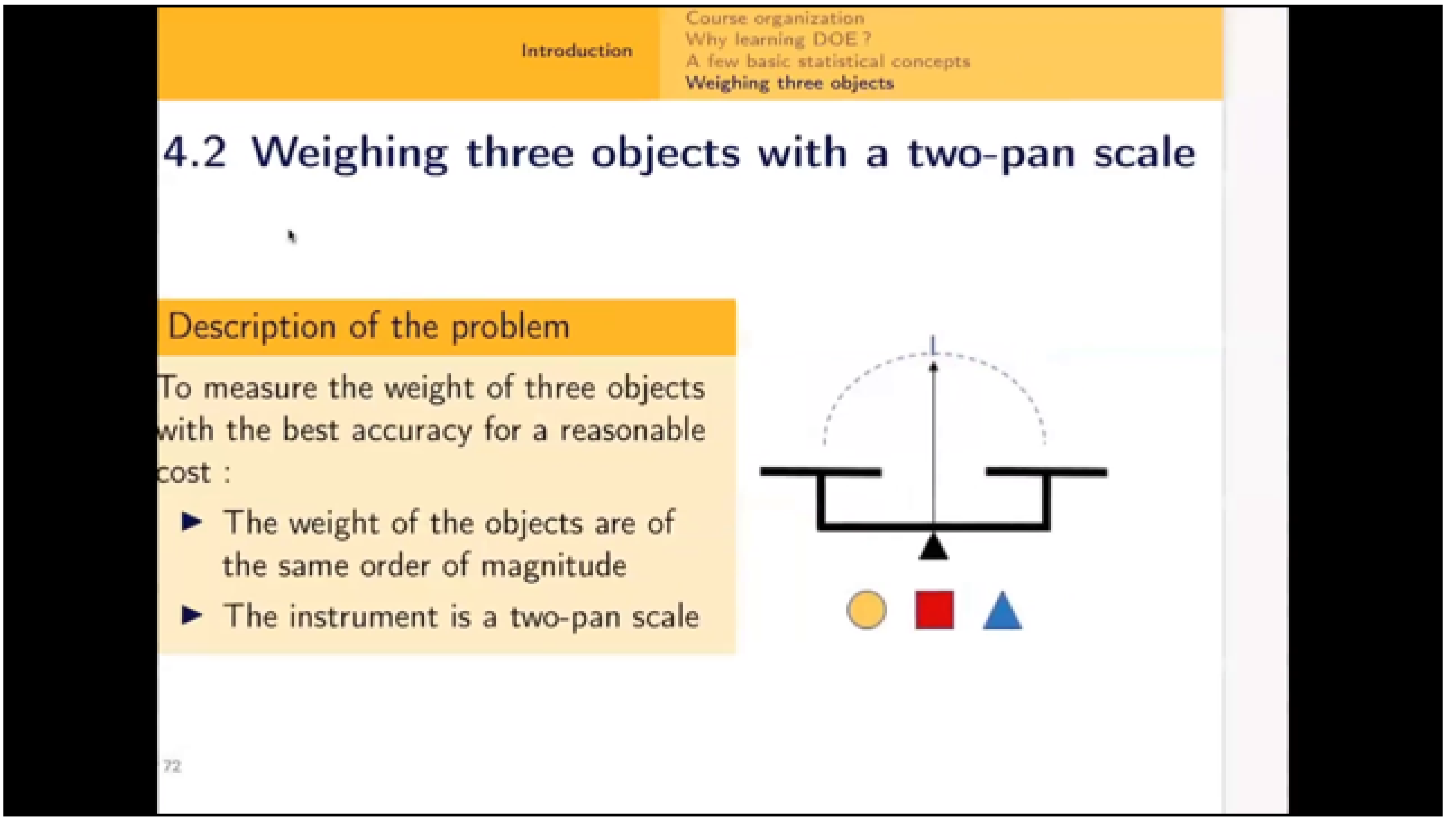


already what is expected in this slide map. Let's see a first case. So imagine that you have to wait three objects, yellow, red one, and blue one. They are all comparable ways, not one at all, and one planet, when you think that you can put, if you want, together in your, in your, in your scale. And in fact, so it's quite a special scale for simplifying. So you can read the weight and the scale, get an equilibrium and unbalance equilibrium, and you can read read the weight here. So you have two pan, and so you can put an object in one pan, and you see

notes

summary

10m 11s

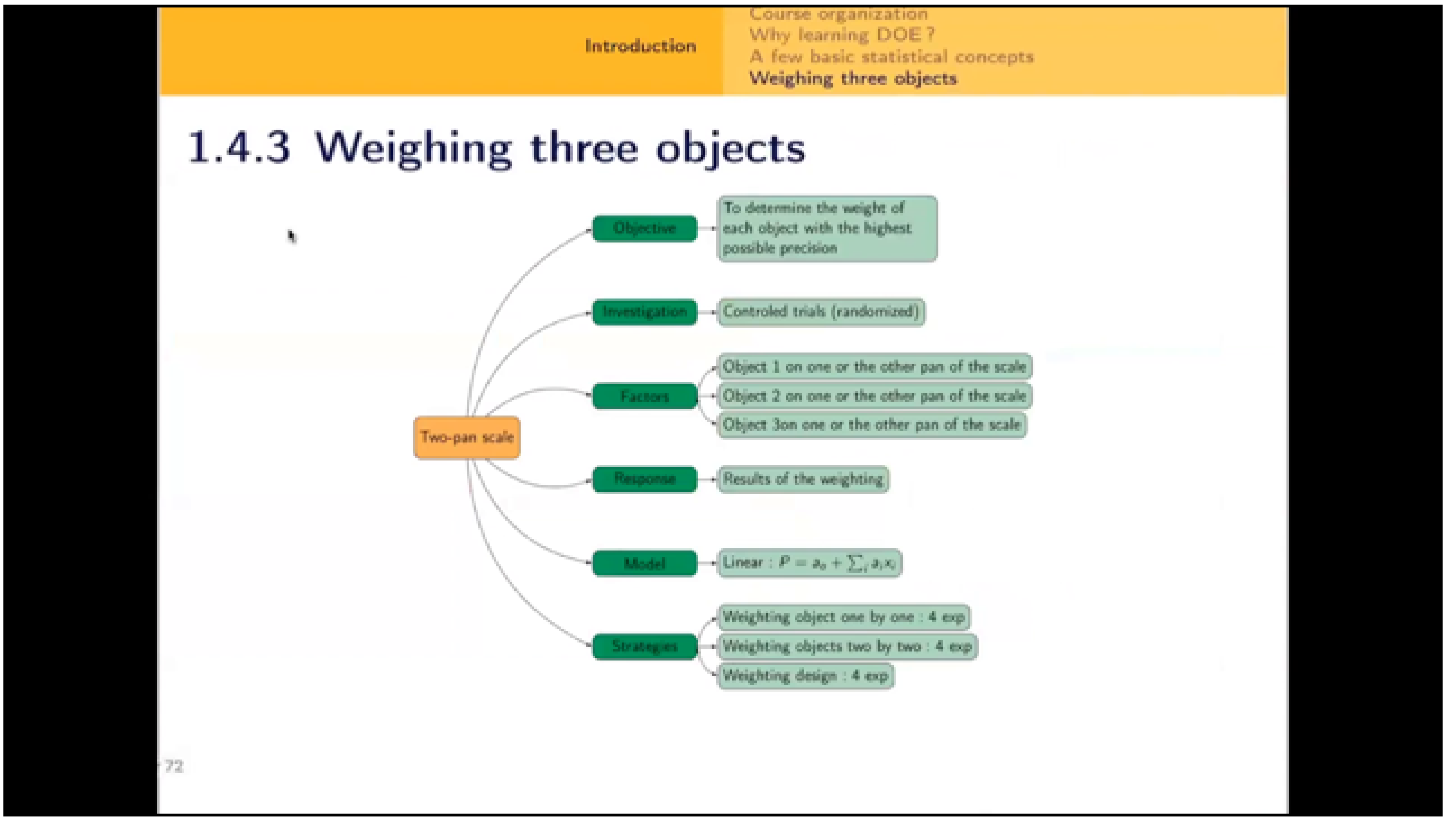


the value of the weight of the object. The idea is to get the three weight with the best possible

notes

summary

11m 12s

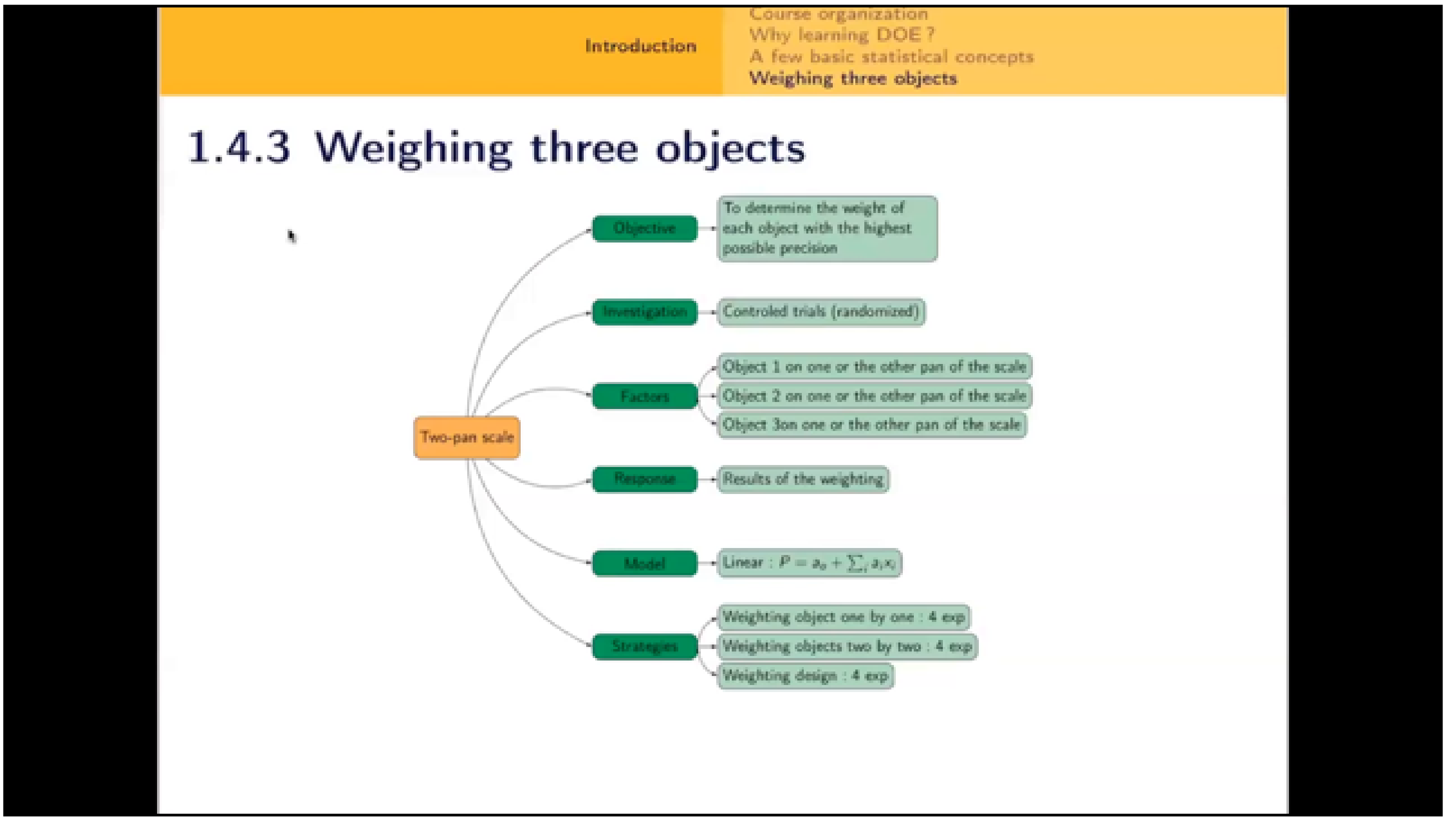


precision. So the mind map looked like that. What is the objective? The objective is to determine

notes

summary

11m 25s

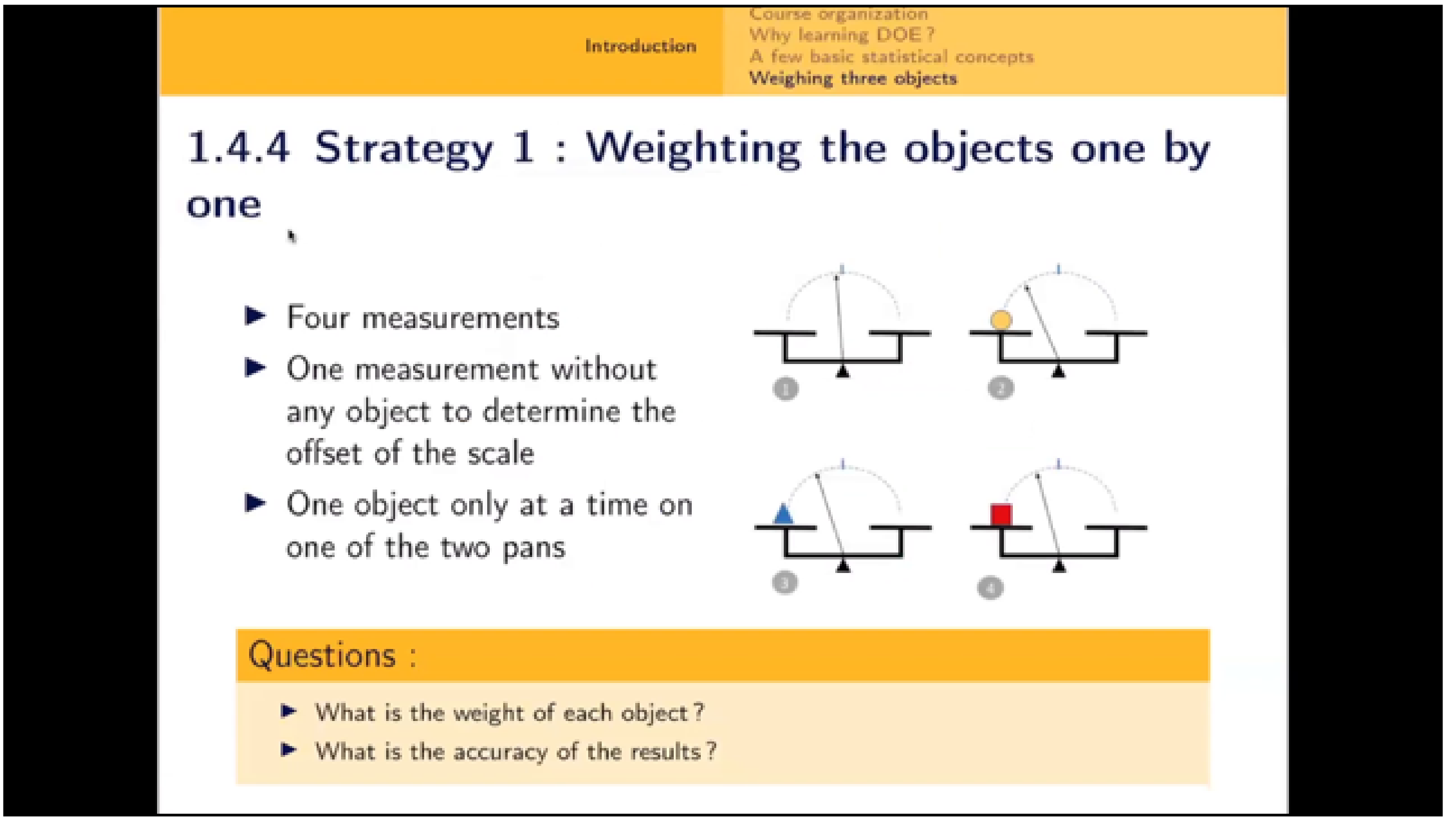


the weight of each object with the highest possible precision. It's an, it's a controlled trial, it's an experiment. I have three factors. So sometimes you can define the factor, it's not very clear. I don't know what to have. Yes, what are the factors of this experiment? In this case, the factor are the object. I'm using the object or not using the object. This could be interesting to think for the master student was because our, our project, you know, are very close to that situation. So one factor is an object one that is a value that the factors can take. I can not put me in the experiment. I can put it in one pan or I can put it on the other pan. It's a different value. I have, in fact, three values for each factor. And I have three objects that are exactly the same. The response, also my experiment is weighting and the response

notes

summary

11m 26s

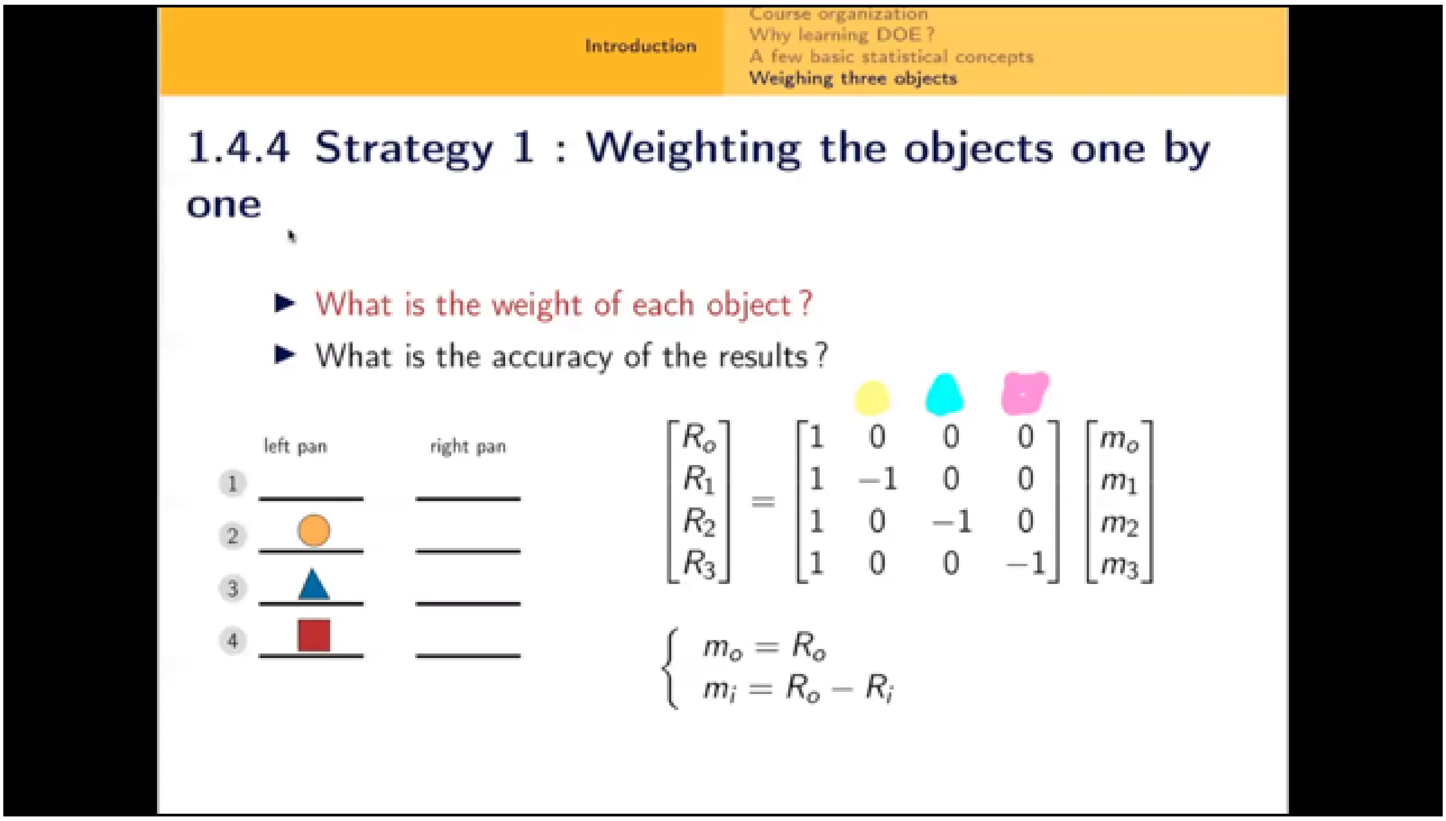


is the weight of what is, in fact, is the difference between the two pan. If you put on the object, it's a weight, but the effect is the weight of my experiment. Of my experiment. I'm reading the position of the needle and it gives me the result. So my model is a linear model. When you put one, two, four objects in the scale, there are no interactions. That simply you add the weight in what we learned from Newton. There are no, in this situation, there are no interactions between the object. You have not to consider because it's a system. So eventually the mass of one object is attracting the mass of the object that will not change the total mass of the system. I will show you different strategies and I will show you the best exercise for this week will be to play with even a small situation, not only a three-object that will make you play with

notes

summary

12m 49s

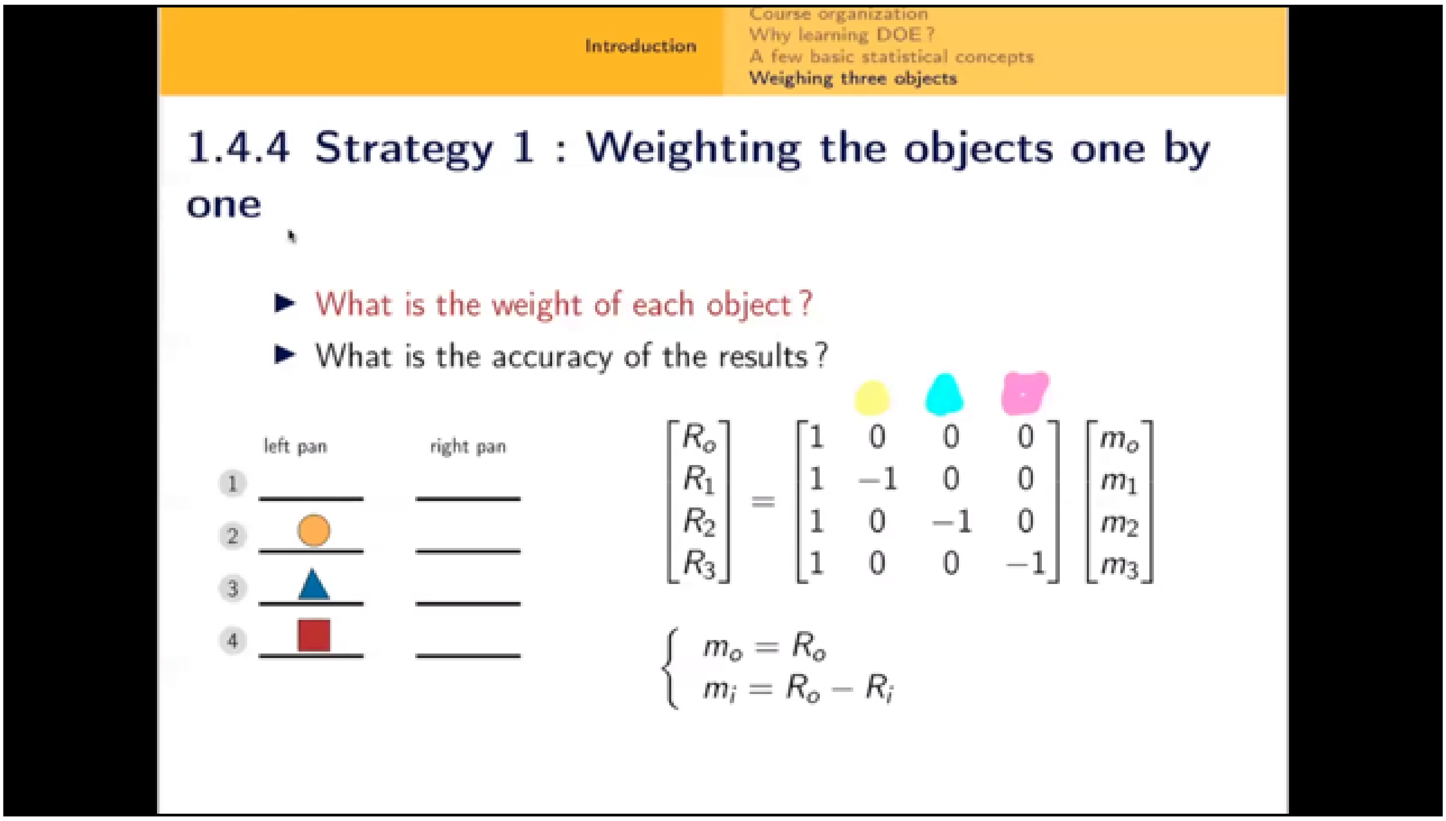


that. The first thing that we have in hand, I didn't ask you to guess, but very profoundly, you check first your instrument. So you made an experiment with nothing for checking the offsets of your measurement is what we have to do. And so you have to know to understand that making the measurement of an offset is an experiment. You have to count them as an experiment. And after I put each object one after the other in one. You could also have a new maker's other if I need to not change our results. So we each time have two questions. What is the weight of each object? It's what I would like. And also we have to define the accuracy of the weight of the object. See how I'm now representing my experiment. I have four experiments and in each experiment I just draw where I put which objects and I can represent that in something I'm calling a matrix of the model. And so my first experiment I have nothing. So if I have an offset, the offset is present, but I have no object on the panel. So each, so the first column represents the constant, the offset that is present all the time. And the other columns represent each object. So when I'm writing zero, that means that I'm not using the other. When I'm putting minus one, that means that I'm putting the object on the left span. I decide, I said, this one, this one will not change nothing in this case. But I'm considering that my needle is indicating negative weight in one sense and positive weight in the other. So my second experiment will be one because my offset is present and minus one because I have put the object on the left span. And the third experiment is one, zero, minus one, zero, because I have my offset.

notes

summary

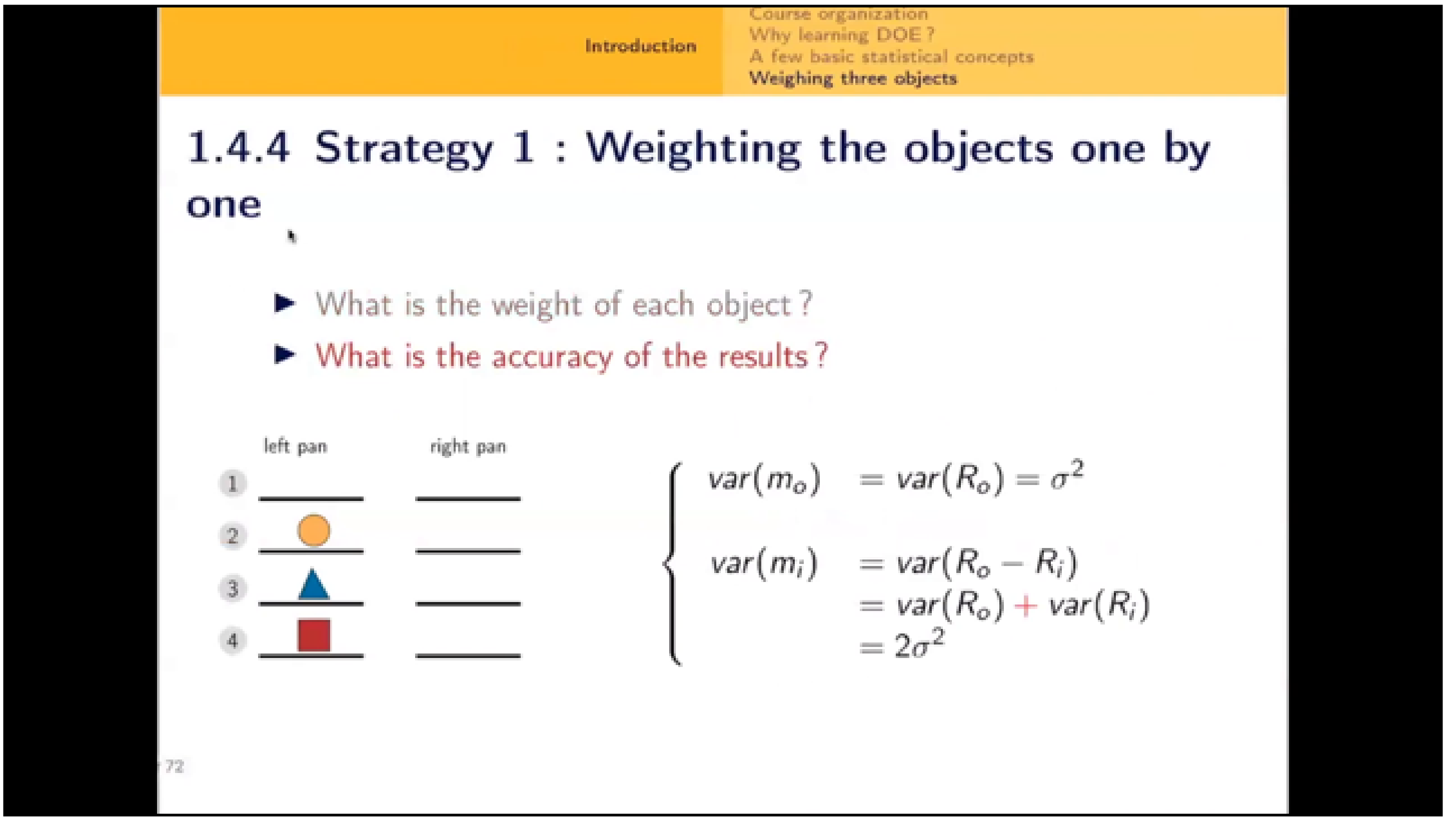
14m 1s



I have done nothing with the yellow object. I have put the blue triangle pyramid on the left span. And I have put nothing with the red object. And the last one will be, for this situation, I don't need to make this complicated situation. But for the next one,

notes

summary

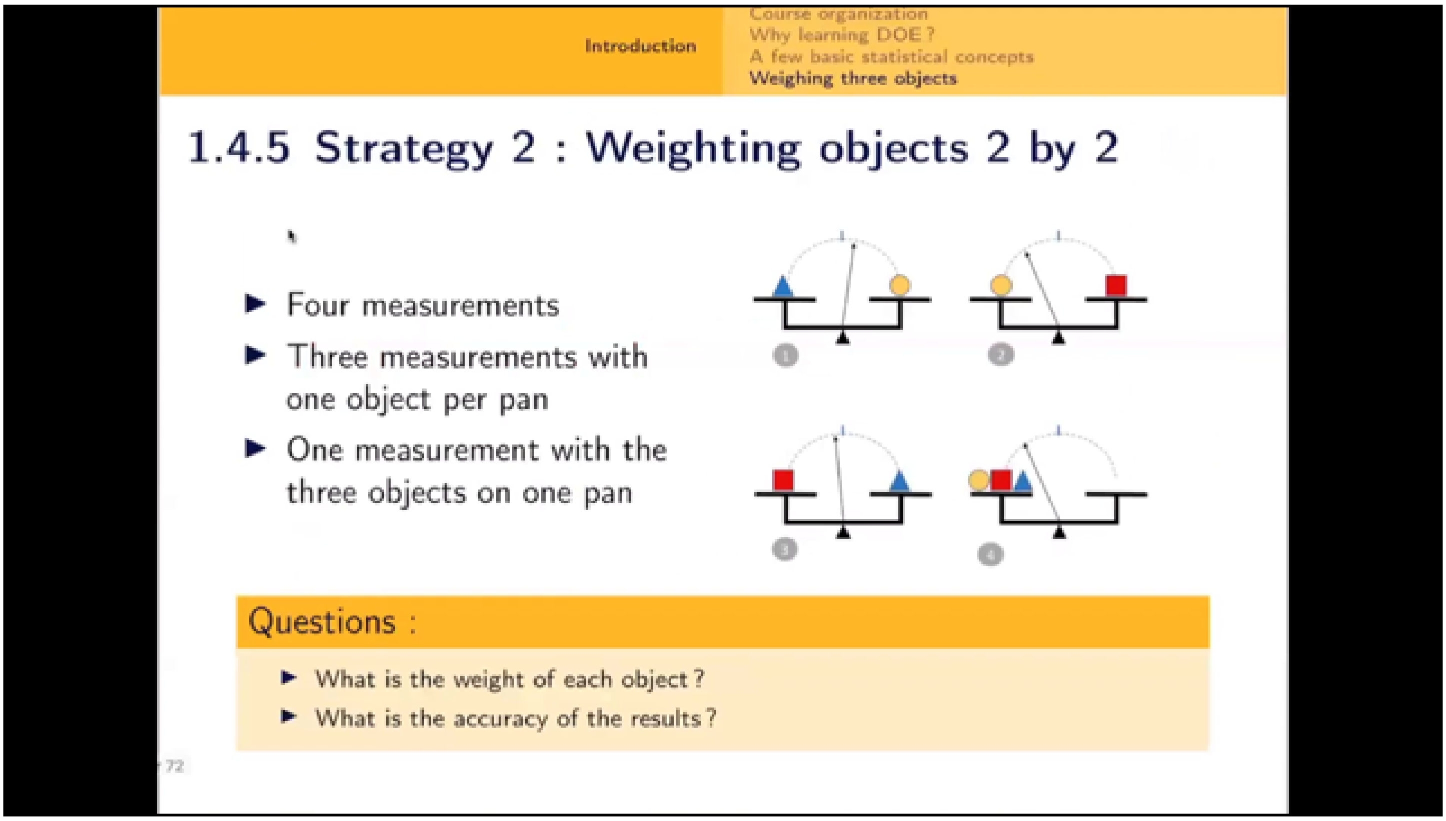


I will need this representation. So this is called matrix of the model. And you see that you have a linear system. I have written here my unknown, the object of each of the, my weight of each of the object. And my offset is m zero, I put m zero is the offset. It will leave the unbalanced situation of my scale when I have low object. So you see that it's constituting a linear system. You can say that the result of my first experiment will be my offset. Because it's a multiplication of this line, this row with this column. And for the second experiment, I call it r one. And it will be the offset minus the object on the, the first, the first object that I have put on the left span. So it make your system of equation and the solution in this case is very easy. So the offset is the first result of the first experiment. And the weight of each of my object is the measurement of the offsets minus the measurement of the result is minus because of the sum. So the second question is where things start to become interesting.

notes

summary

17m 1s

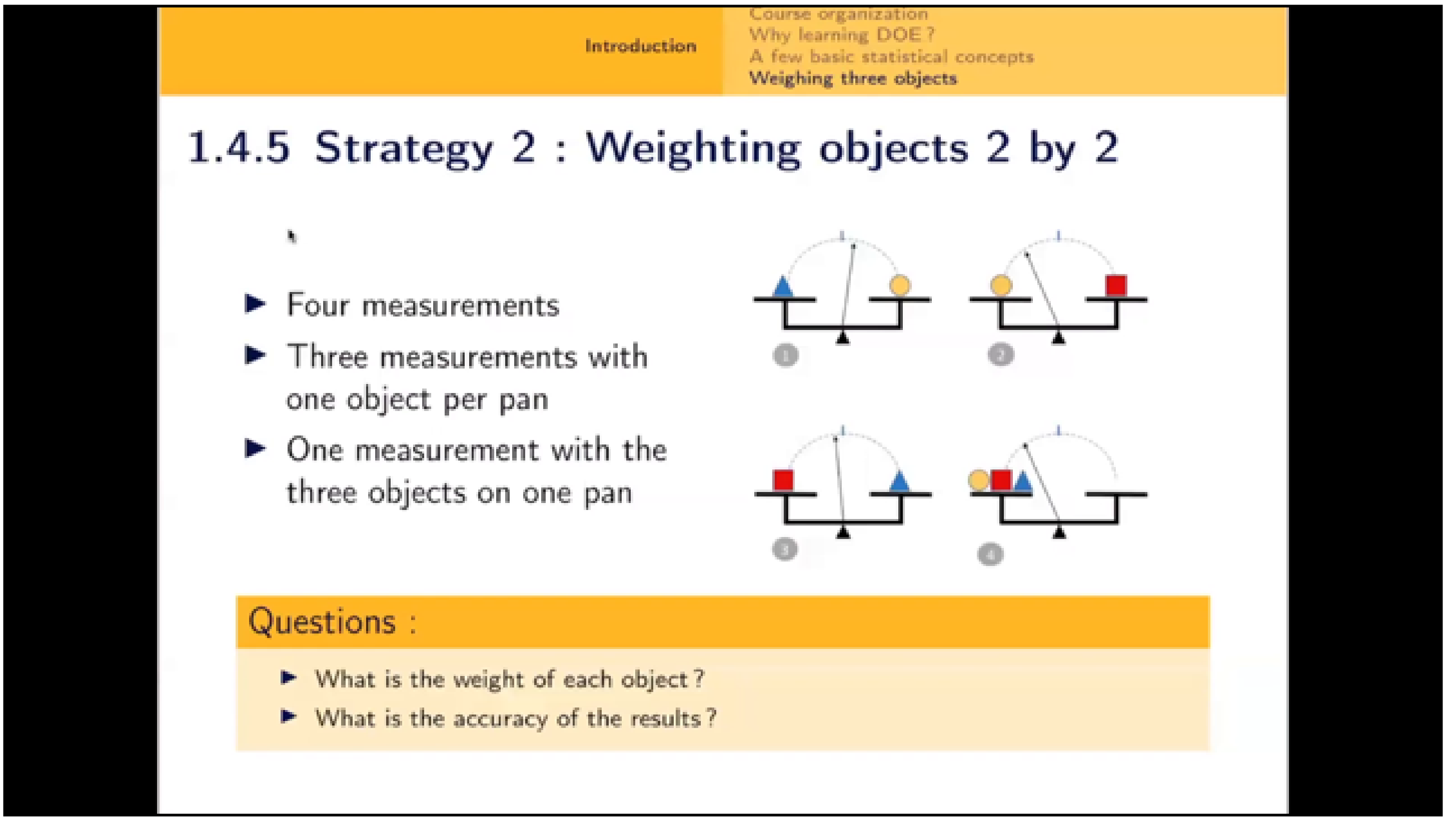


Well, the variance of one of the weights is the variance of the estimator we are using. So for the first one, the variance of the offsets is the variance of the first experiment. Let's imagine that sigma square is representing the accuracy of my scale. That means that I'm able to make a lot of measurement in my scale, understanding the variance of my, of my measurement. So it's clear that it's, I know very well the offset is not so interesting, but in any case, I know the offset very well with the precision of sigma square, variance of sigma square, then I'm able to calculate the confidence, etc. But now when I'm interesting by the variance of each of the weights of the weight of each of my objects, well, I have the estimator is the difference. And you know what we call the law of variance, the law of variance, that's the variance of the difference is the sum of the variance as returns. So that means that variance of r, 0 minus r, y is the variance of r, 0 plus the variance of r, i, when r, i is the result along to all three, that means that the accuracy of my measurement of each of my objects is in fact, true time, the variance of one measurement. It's not a good news for you, but I have a solution. That is something we forget because usually we forget to talk about the offset.

notes

summary

18m 45s

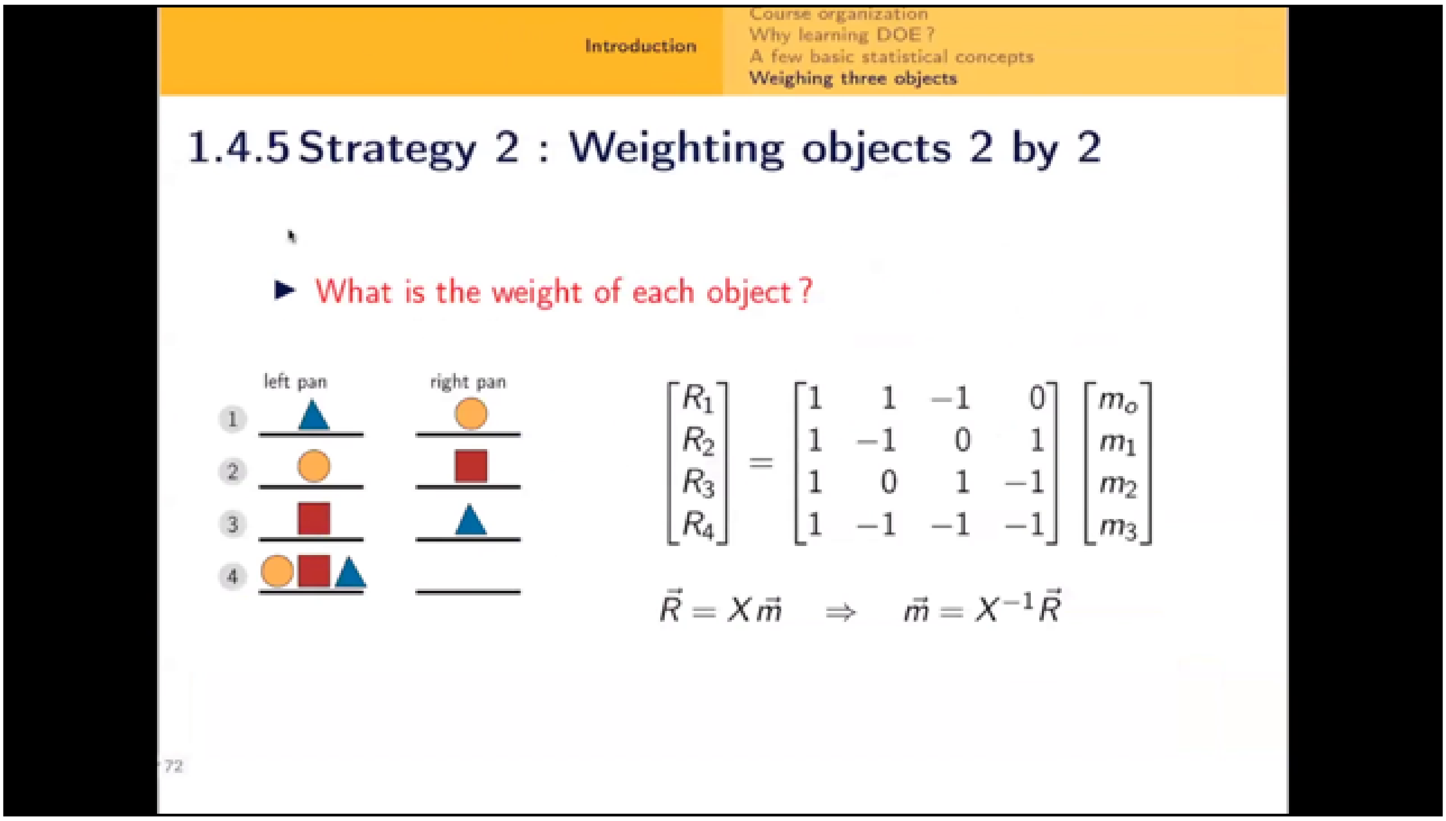


So now let's try to make something different. So if we do something different, we can try to put not one object at the time, but two objects, one on one side and one on the other side and making linear system, which have a rank, a sufficient rank for solving it. So you see I put

notes

summary

20m 42s



the blue and the yellow, the blue and the red, the red and the blue. And after if I want to solve the system and having a rank, the fourth experiment, I will put all the objects on one side, doesn't matter which one, if you don't do this experiment, the system is going to have an acceptable rank.

notes

summary

21m 13s

1.4.5 Strategy 2 : Weighting objects 2 by 2

- What is the accuracy of the results ?
- $$\begin{bmatrix} m_o \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 3 & 3 & 3 & 0 \\ 4 & -2 & 1 & -3 \\ -2 & 1 & 4 & -3 \\ 1 & 4 & -2 & -3 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{bmatrix}$$
- $var(\vec{m}) = var(X^{-1}\vec{R}) = (X^T X)^{-1} var(\vec{R}) = D \; var(\vec{R})$
- $var(m_i) \approx D_{ii} \; \sigma^2$
- $D_{00} = 1/3 \quad D_{11} = D_{22} = D_{33} = 10/27$

So what is the weight of each object? So now the system is the same structure as before, the values are different. When you have an object on the one for the constant of the time here, because the offset doesn't change during your experiment, so one, one, one, one, but after the yellow object, I put it one time on right, one time on left, one time was not present and one time on left. And the same thing for the blue one and the same thing for the red one. And so I have a linear system and the rank of this linear system is sufficient for inversing my matrix and then the object of each of my, the weight of each of my object, even the weight of the offset is the inverse of the matrix X, with the matrix X is this one. Multiply by your vector of results. Nothing very special. Now let's talk about the accuracy. So what is the accuracy of the results? So if you have a linear system like this one, I have put some some value. So you have your object, which is the inverse of the matrix. I've calculated the inverse on the X matrix, one line of the few elements multiplied by the vector of R. So the variance of the vector of weight will be the variance of the estimator. So I don't know if you are used to make calculation of variance with matrices, but you know that the variance function is a quadratic function. So when it's not linear, it's quadratic. So if you have to put something in evidence, you have to take it this way. So you need to take it. So the X, my experiment is considered as a not factor factor experiment, but a factor like the alpha and multiplying X. And so if I want to put X minus one

notes

summary

21m 35s

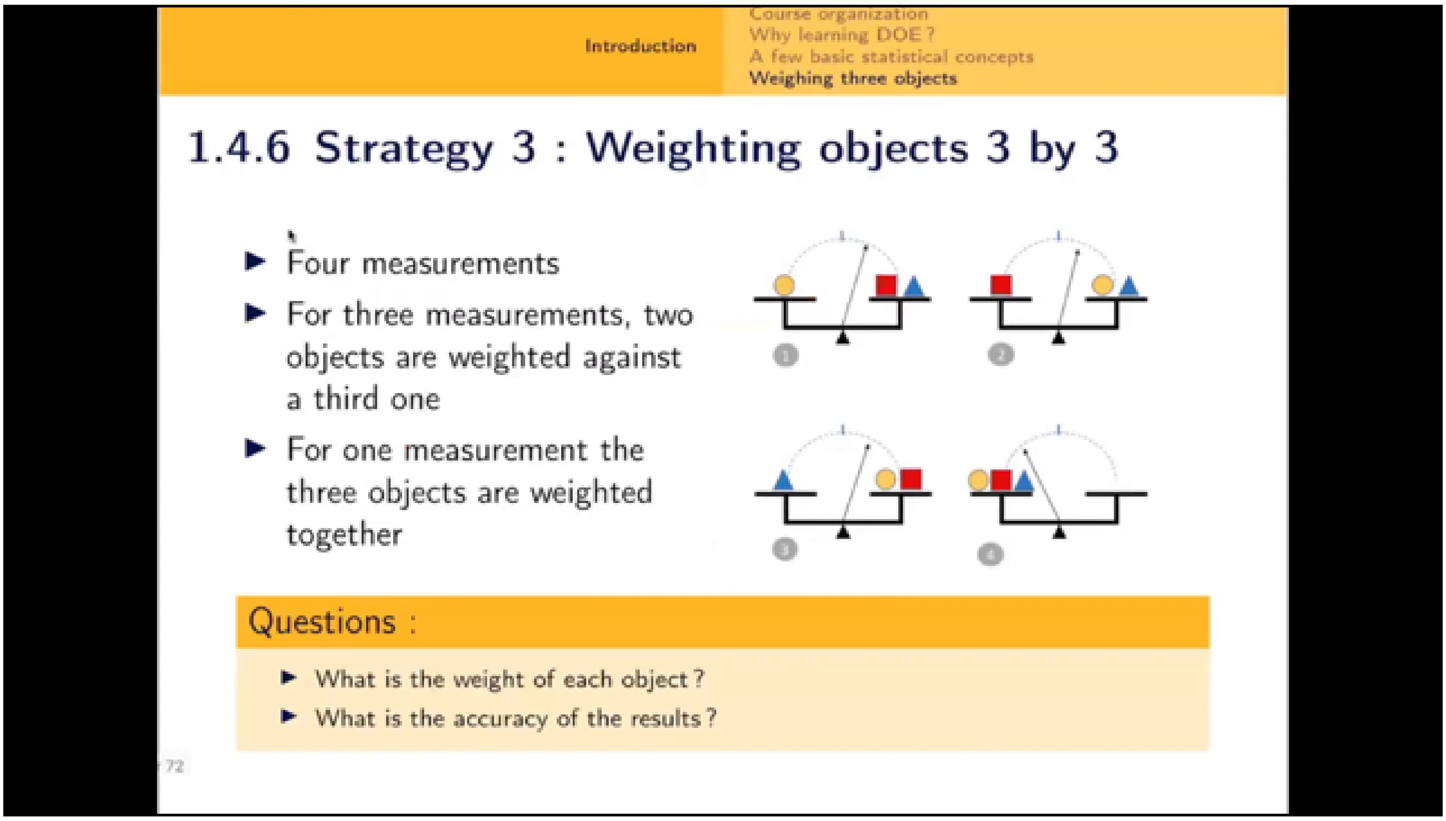
1.4.5 Strategy 2 : Weighting objects 2 by 2

- What is the accuracy of the results ?
- $$\begin{bmatrix} m_o \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 3 & 3 & 3 & 0 \\ 4 & -2 & 1 & -3 \\ -2 & 1 & 4 & -3 \\ 1 & 4 & -2 & -3 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{bmatrix}$$
- $var(\vec{m}) = var(X^{-1}\vec{R}) = (X^T X)^{-1} var(\vec{R}) = D\ var(\vec{R})$
- $var(m_i) \approx D_{ii}\ \sigma^2$
- $D_{00} = 1/3 \quad D_{11} = D_{22} = D_{33} = 10/27$

in evidence, in fact, putting in evidence, taking it square, because of the matrix calculation, it's became X transpose X minus one. You have to believe me or to check in your algebra. So now the variance of M is the square of the coefficient putting in evidence, X transpose X minus one, multiplied by the variance of R, which is an experiment of random variable. X is not a random variable. When we are making linear regression, in fact, we are considering that the X values are defined precisely. We're not considering that uncertainty there. We're considering uncertainty is a risk of Y. It's not good, but it's not. So that means that if I call this matrix D, we'll have a name, next week I will see all the structure of that. It's a dispersion matrix multiplied by the variance of R. And we have said that the variance of R is the variance of my scale, so it's sigma squared. So finally, you see that the variance of all the object, it will simplify a little bit. We consider the matrix as diagonal. We can consider that the

notes

summary

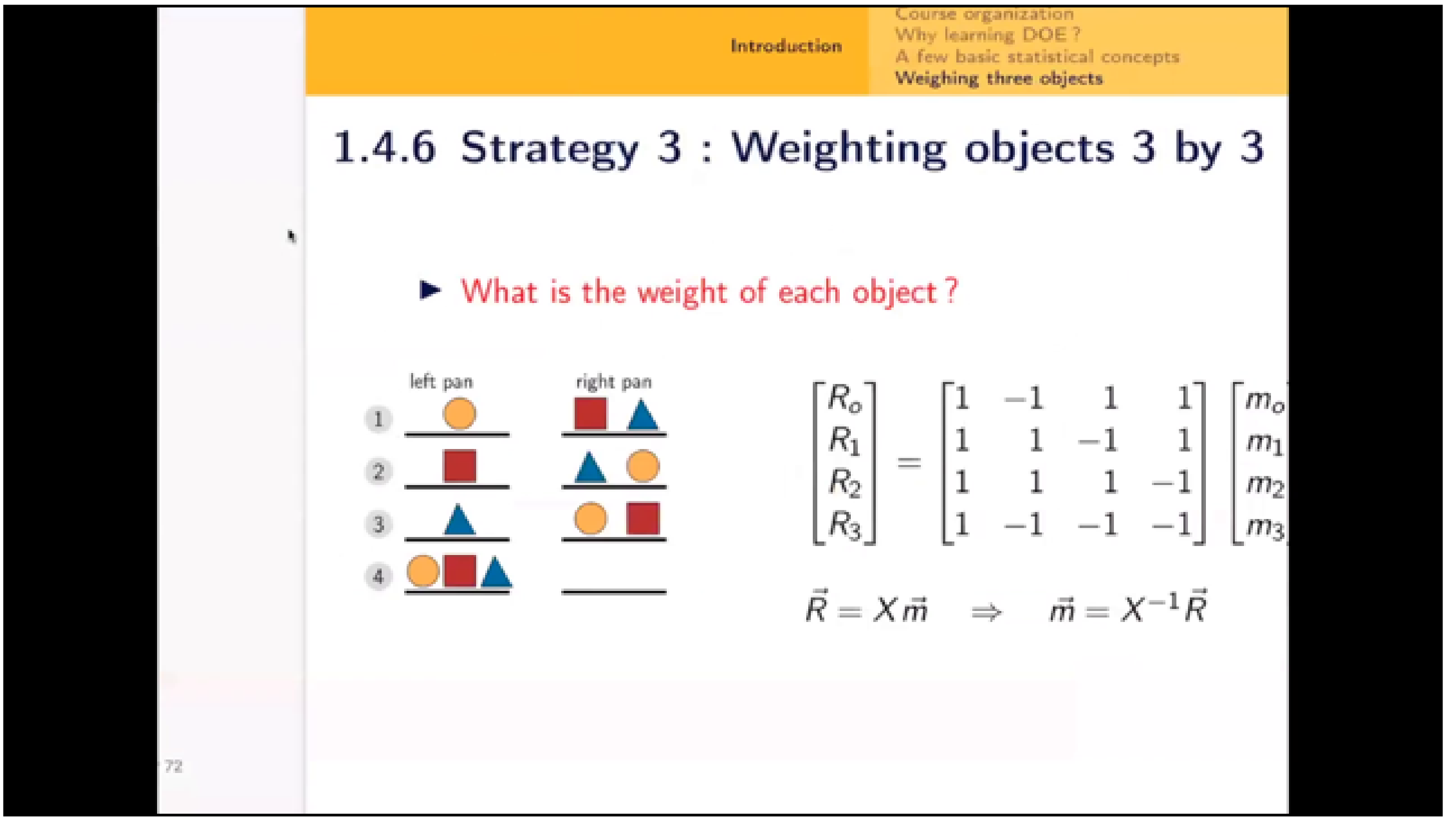


variance of each of the weights is the diagonal element of my matrix D multiplied by the variance of my method. And if you make the calculation, you see that the variance for the offset is one-third and the variance for the three objects are quite close to one-third is one divided by 27. It's a little bit better. Before you remember, it was true. No, it's one-third. So I have improved by a factor of six. It's the same cost, four experiments. Okay, I have to move a little bit more your weight. Is there a time, perhaps, you're moving the frame or moving them, perhaps you could move a little bit more. But in fact, these four experiments are quite the same cost. And now we have multiplied by six the quality of our root of six because it's a variance, which we call just a confidence interval. So root of six is something I'm not going to do in three or something like that. So I improve my quality quite more than two times just by changing my strategy. This is design of experiments. This is what we look for. But I can even sell you something better. It's more expensive, but it's better. No, it's the same cost.

notes

summary

25m 37s

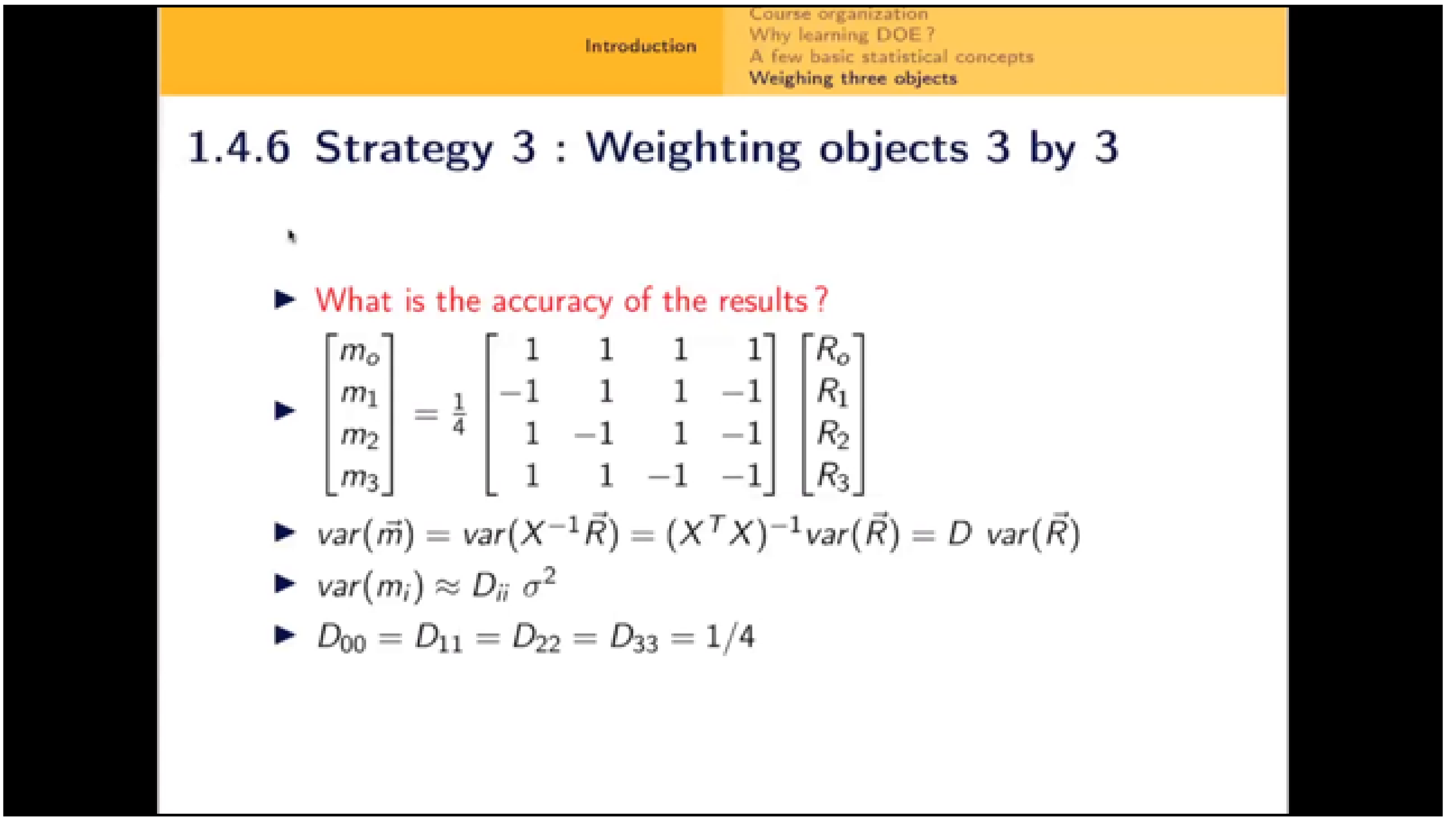


Now let's try to do that. Each time I'm uniting all the objects, the yellow one at left, the blue and the red one at right. And after I turn around and my last experiment is the same as before. I needed for having run sufficient in my system. So you can see that in fact the left span FC six objects and the right span FC six objective. One, two, three, four, five, six, and one, two, three, four, five, six. It's a balanced design.

notes

summary

27m 0s

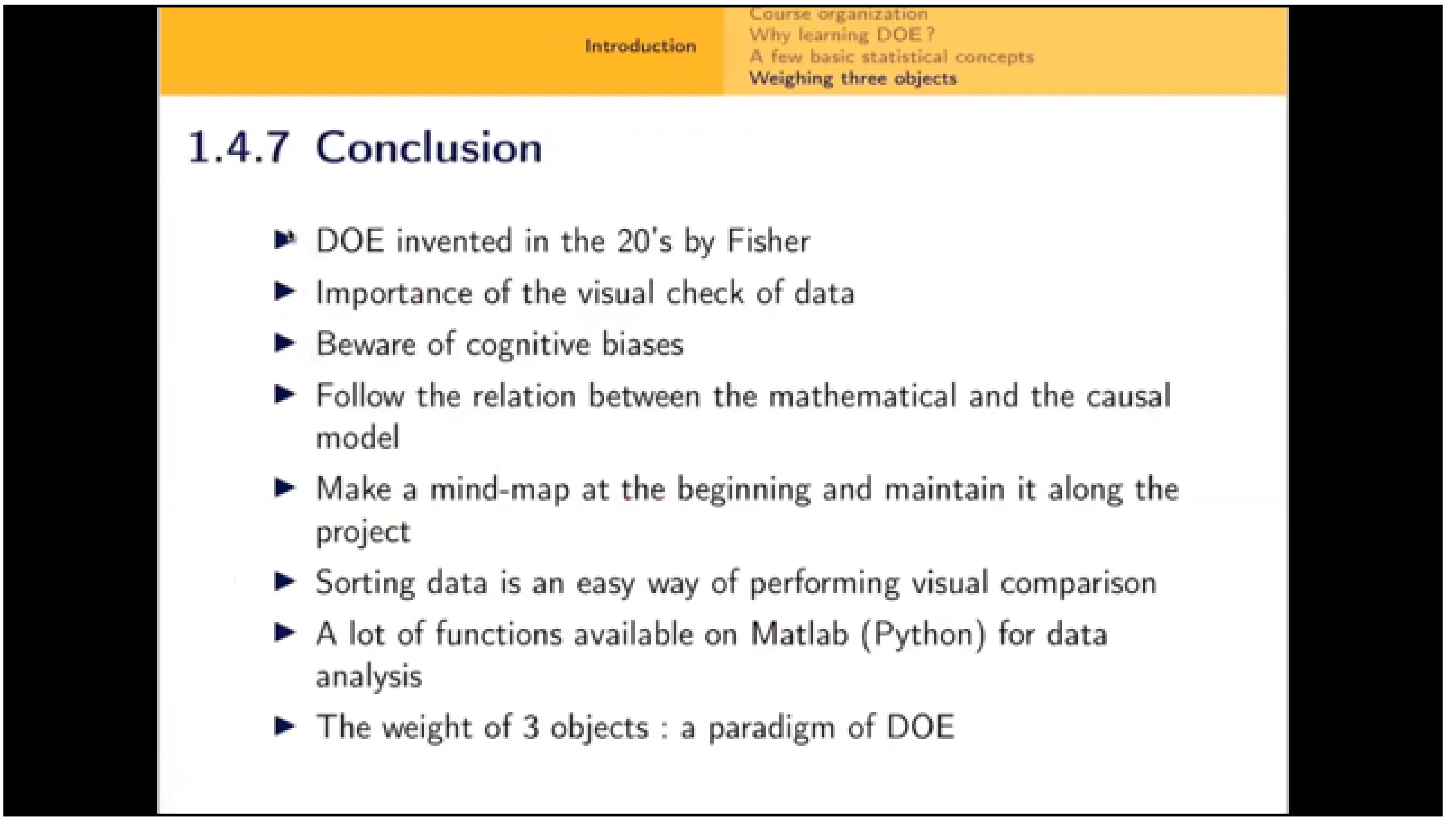


If I want to see the values of my object is as before I have a linear system. I can solve it. And then my weights are x minus one, the inverse of the matrix x, the matrix x is this one, the matrix of the model. And I buy my vector of my vector of results. But now if we do the same calculation as before, was this matrix. So this is the inverse is a yes, is the inverse of the matrix x one, the force of some ones, but sometimes plus sometimes minus. If you do this calculation, so the two first line are the same as in the previous slide, you do this calculation. Now you have the same quality for all the values. So my my offset is calculated with the same precision as the father. And my quality is one force. So between two and one force, I have improved eight times my variance. So roots of eight, two points, seven to something. So I have improved my quality quite street time for the same cost for experiment. This is what we look for. We look for not making more experiments, but having a higher point in your in your output. This type of

notes

summary

27m 56s

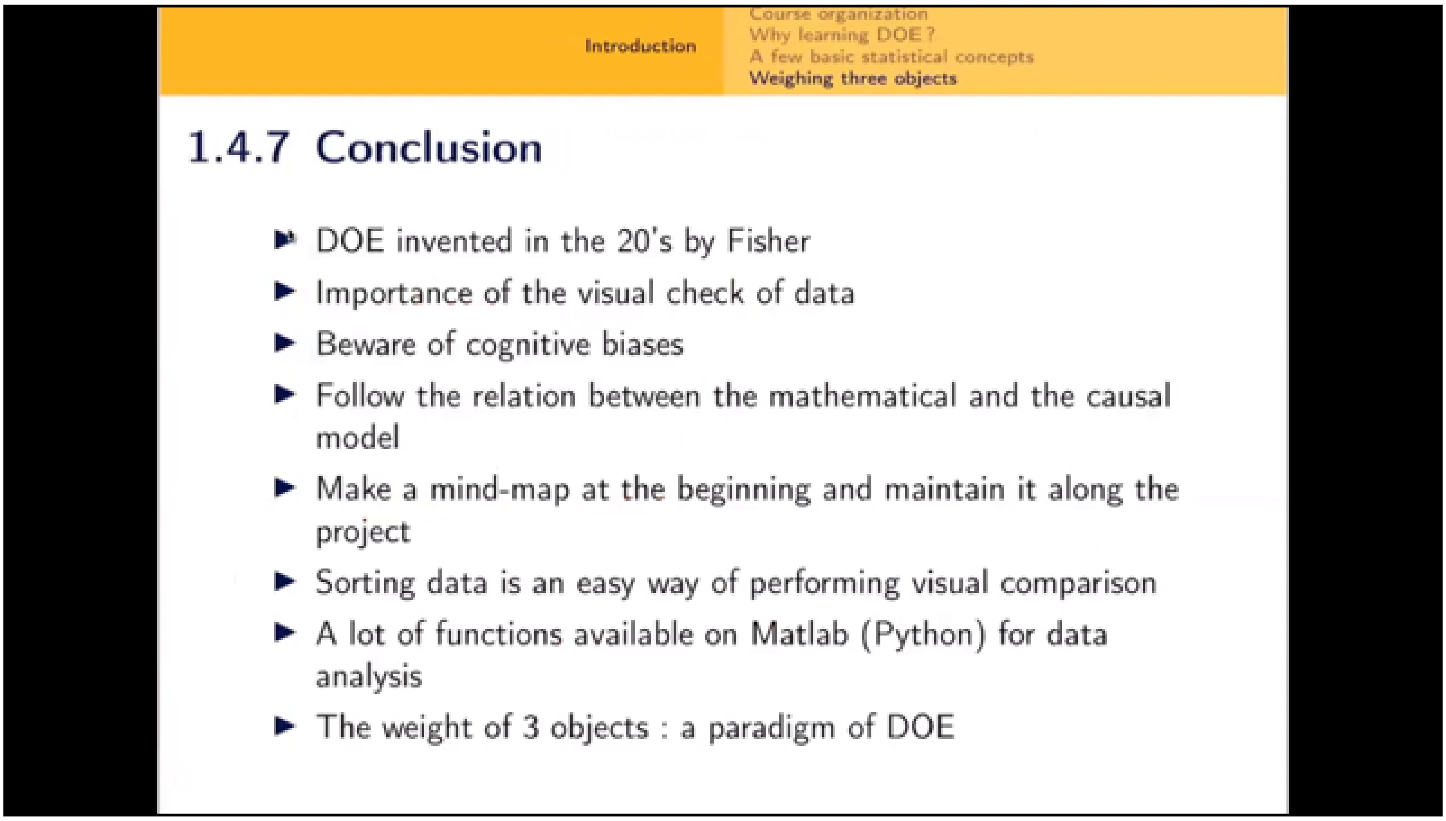


matrices, have a name we will we will learn them. This is called Adam or matrix is a matrix that have a property that when you multiply by its transpose to make a diagonal matrix. And so it have a lot of advantage in math lab, you have a function called Adam or and the number this one will be an Adam or four, the numbers of run, and it give you for all the multiple of four, all the other matrices. It's not the best solution for everything, but for this situation, for waiting factors when they do not have interaction between them is the best possible. They are also theorem showing that it doesn't exist a better than that. You can make more you can make measurement more times for having better quality, but in the structure of the experiment is the best possible. So as a summary of this introduction, DOE have been invented in the 20th century by Fisher and the 20th century. of the visual check is very important. I try to make you attentive to the fact that we have bias or so be careful with your bias. We have a talk about this causal relation. It could be more, it could be more description, it could be of course about logic, and about epistemology, just to make you sensitive to that, but I didn't spend too much time on it. After I present you the interest of making a mind map of the situation for having with six, seven information, what is statistically speaking, what do you have from the list? If you have made your series one, you have this corner, so that the salty new data is something that can also start to take out information, showing information is very important when manipulating data. There are a lot of functions available on Matlab and Python and R for this type of

notes

summary

29m 49s



operation or manipulation, so take time also to read the help, understanding those functions. I have put last week on model some chip page about Matlab when you can see a few of those functions. And today I have show you how to wait three of Jack. I hope testing for you, you will see exactly what we are looking for, what we will be doing, so the rest of this course will be making the correct calculation because of very simple situation. So generalize the calculation we have done. So next chapter is about modeling, how we will model the situation and after we will look for some very typical classical design, the placket and remanence will be Adam or Mattresses that are used for making placket and reman designs, the full factorial design, the fractional factorial design, after we will go for people's surface responses,

notes

summary