Course material

Course:

ENG606 / PHYS 442

Video:

DOE_lesson1_part3

Concepts (extracted from automatically generated subtitles):

Type of data. Quantitative data. Very important health problem. Only thing. Group of students. Much problem. Okay pareto principle. First cookie. Second world war. Obscure important variation. Product b. United states. Highest value. Percent rule. Children of a very tall person.

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

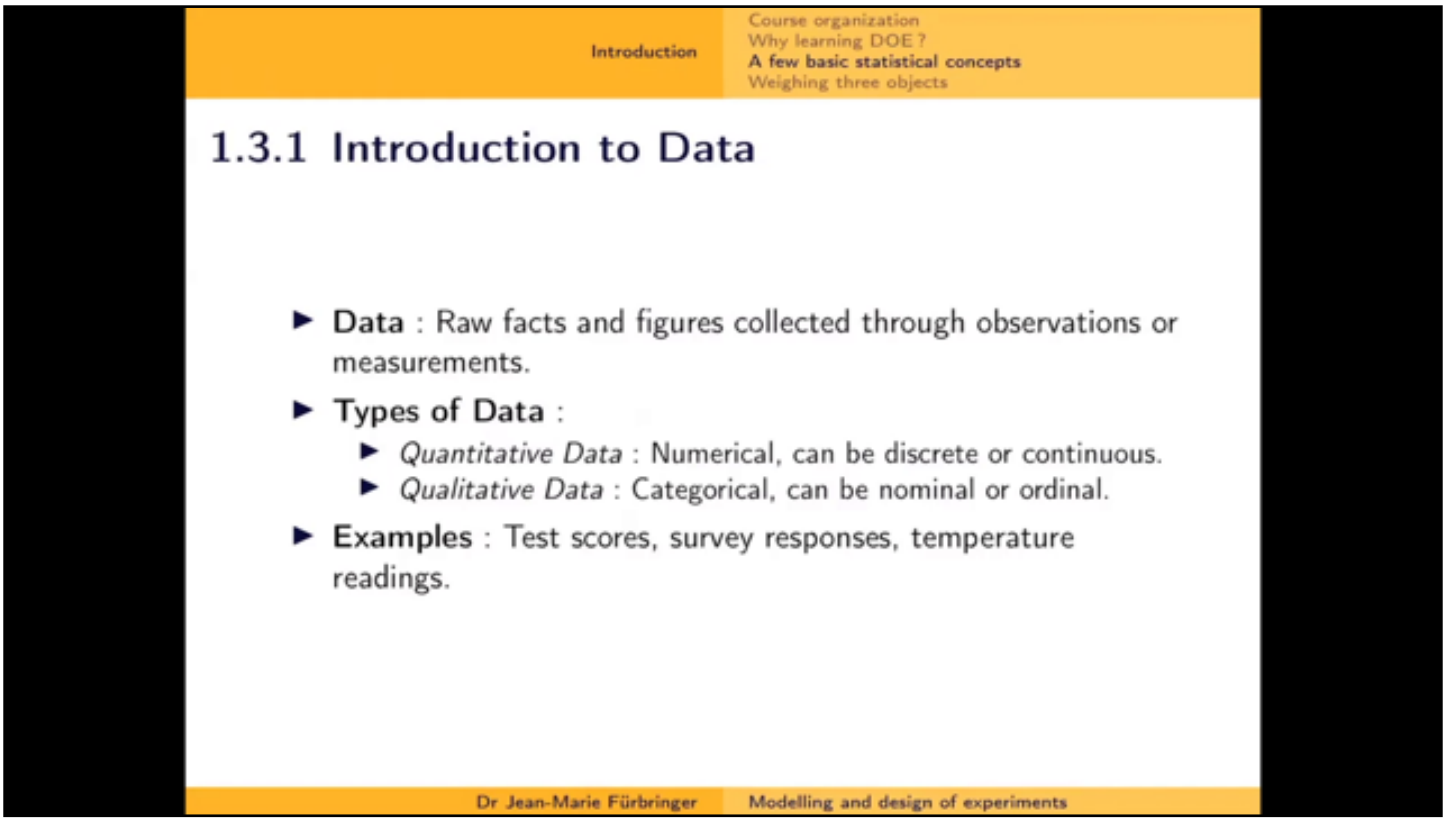
to video

...

notes

summary

0m 0s

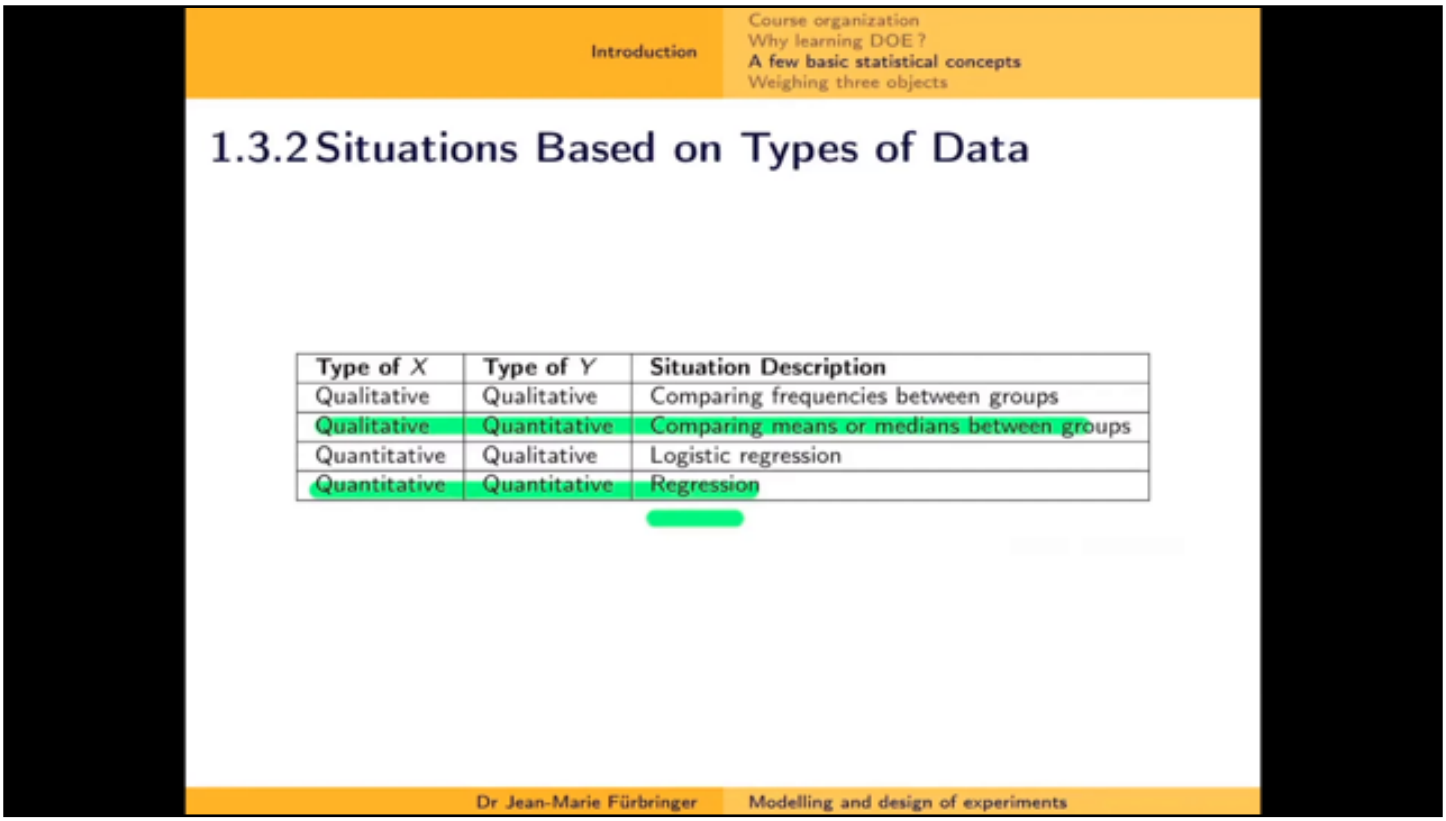


These subtitles have been generated automatically

notes

summary

0m 1s

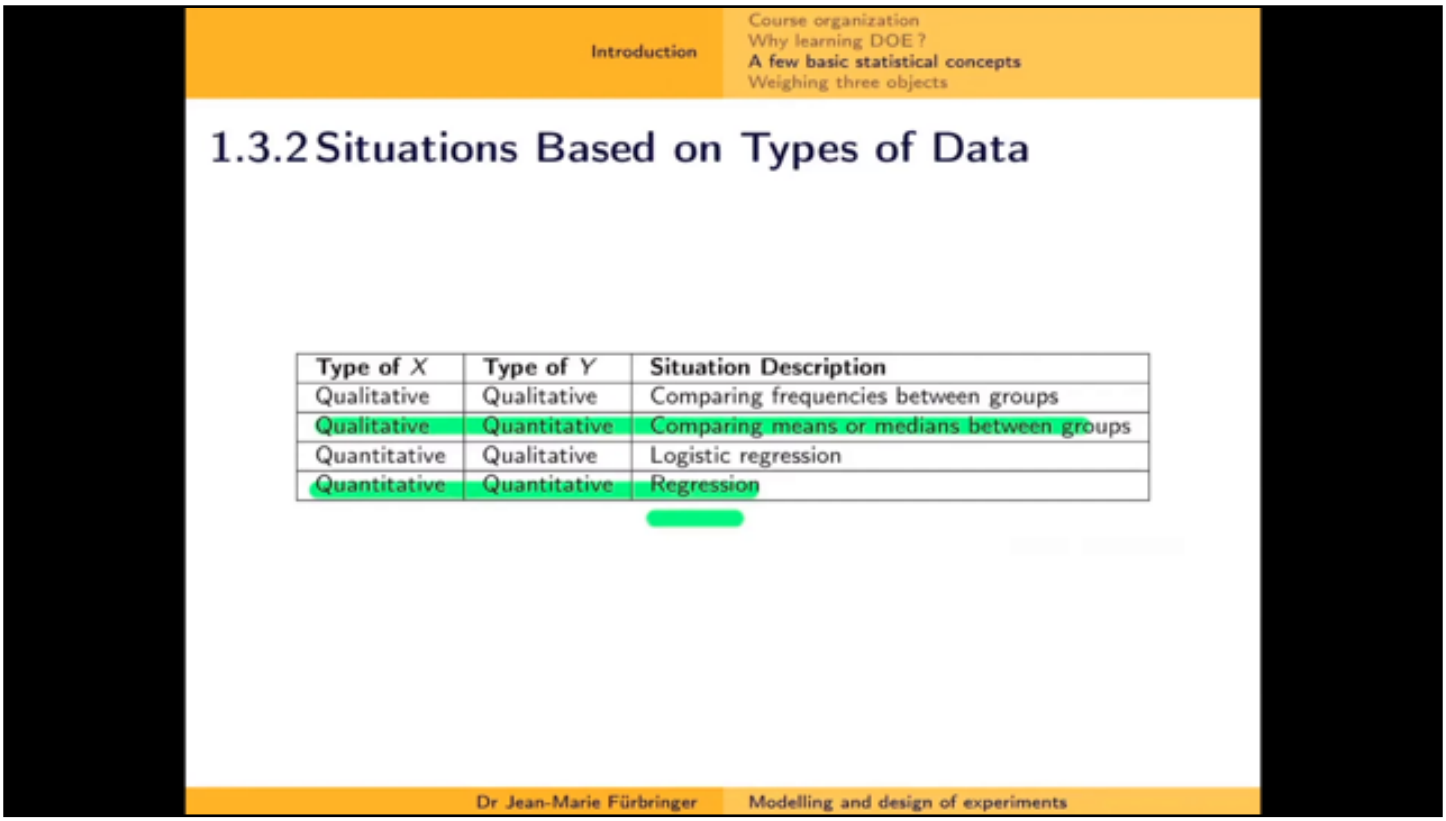


I asked yesterday, chatGPP to make this slide. So I said, OK, I had one slide quite obscure. I was not happy for many years. And I said, chatGPP. So the next slide are from chatGPP. They should write chatGPP. OK, so type of data is something very, very, very important to understand. You have quantitative and quantitative data. And they are not treated the same way. We cannot do the same. So what we call quantitative data. So it's a value, 5, 3, 4, 1,000, 2 millions, et cetera. And when we do qualitative, so it's a category. So it was the product A or the product B or the product C. We first cookie. Cookie. Cookie. Or we first dry it. So it's going to be operations. So it's a do not represent by number. You represent them by your category. When it's ordinal so that you can make an order between them, it is quite a little bit the question. You can sometimes put something. You can transfer them as quite quantitative order, and became quantitative. There's a quantity, but a quantity. So type of data, desk, course, survey, response, temperature readings, et cetera. So this are the possibilities. So you have your factor and you have your answer. So we usually use X for factors and Y for the responses, or the responses, just the usual choice. So what you can do. So if you have balls that are qualitative, what's the only thing that you can try to do is to compare the frequency between groups. And after they are statistic, but I will not do that in this question. When they are qualitative and quantitative, you can compare means, millions, and both. We will do that. We'll make something called constant post-vision model. I've been showing a way of treating this. It's happened, but all happens, but they say, qualitative, qualitative in engineering. Why

notes

summary

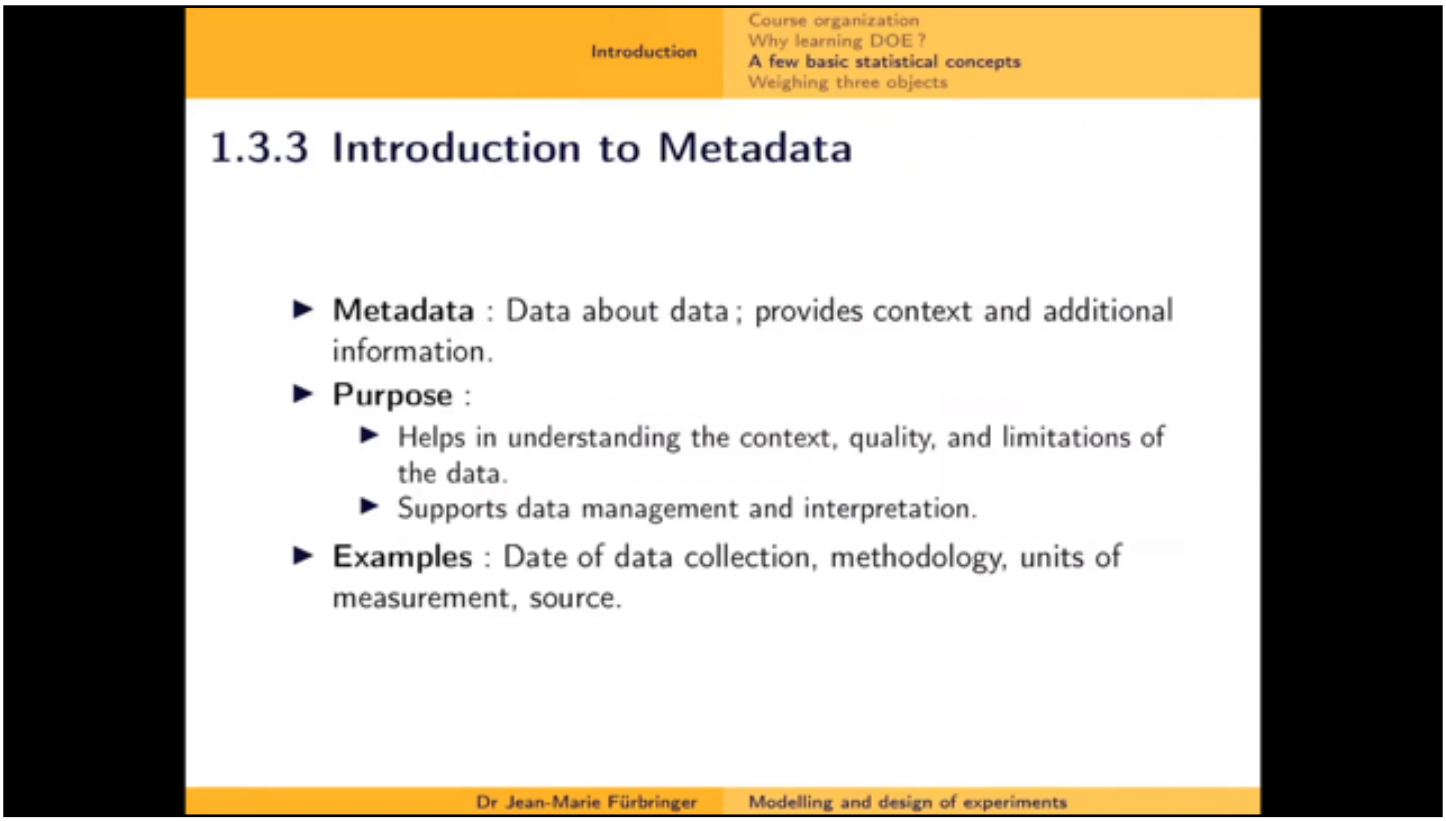
0m 5s



don't I think so much? So we try to usually to avoid, even because it's not science. It's also science, but sometimes complicated science. Quantitative, quantitative. Sometimes people call it, I know that I don't appreciate calling that constant post-mission or random post-mission model. Like that, you want to understand if it's better to attend the course, make the exercise, make reading books in the library. So we can take a group of students and make tests on the results. So it's possible to work with. When it's quantitative, qualitative, typically, it makes some value for your inputs and the answer is fail or not. So we use to make a logistic regression. It's possible for the work. The regulation, the name doesn't... But there are many situations in science where we work with. And watch our course will be most of the time when we have quantitative and quantitative. So you put some value, not more, some resistance and some inductance in your electrical network and you are playing in some position, attention, and current. So you have numbers on one side and numbers on the other side. So in this course, we will work mainly on... Let's say we will work a lot with that and we will work a little bit... With that in the last chapter. I didn't resist to the word regression. I don't know if you know where it's coming. Word regression. It's Galton who has coined this term regression. And its interest was to discover what's happened. Galton is known for something we call regression to the mean. What's happened was the children of a very tall person.

notes

summary

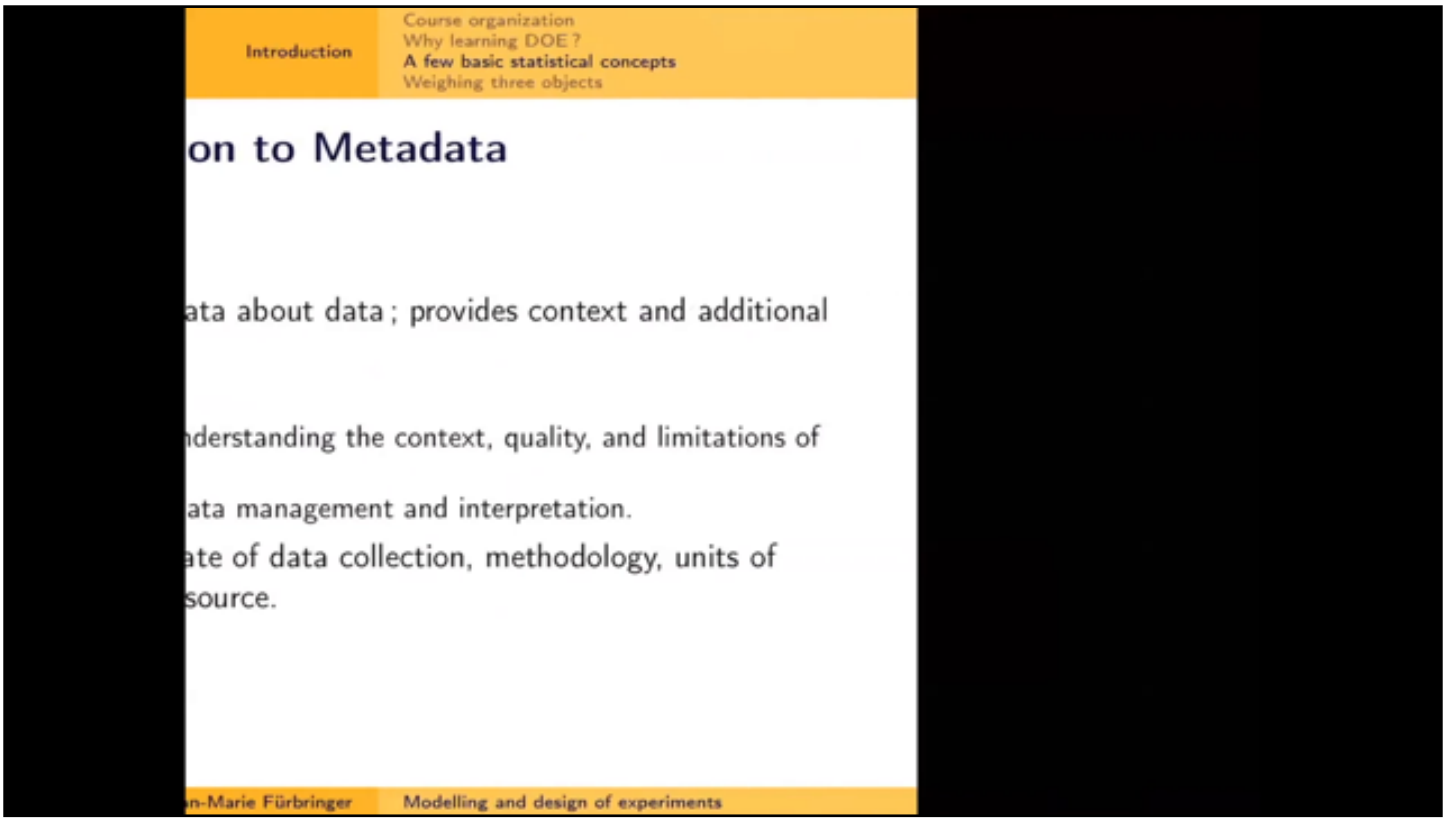


because finally children of tall people sometimes are not as tall as their parents and so the words of repression. I tell you, those three persons have really invented quite everything.

notes

summary

5m 37s



After when we have data, we have metadata, it's also very important. So metadata, the data, the data, the data. They tell you when, where, exactly in which configuration, who and the measurement that we've done. It's very important to keep a straight relation between the data and the metadata. It will help you to understand and sometimes it will make a big change to analyze the wrong data. Maybe because what you have in mind, but not only the conclusion would be completely wrong if you don't take it. So I really insist on the necessity to keep a very straight link between data and the data. It's also why I recommend you to use in MATLAB more tables and arrays for your data. Or when you make Excel files, a lot of people keep the data in Excel files. It's very, very important to keep track. And another point, laboratory notebook, I see students taking notes. This is the metadata portion. But for the metadata in an experiment, use a laboratory notebook. Unfortunately, in this school, I don't understand why it's not a culture. Now, some laboratory insists obliged that in the general culture of EPFL, unfortunately, it's not. We didn't learn that students in TPP are not used to make a laboratory notebook. It's very important because it's also a good way to be sure that this file is the file you have taken. When is that, the fire clock, when we have this, I don't know, the window was open or it was sunny or the temperature of the laboratory was very low. It really lets you to understand what you are. So very, very important. The unit is also the stabilization of your device and anything which was the device that you use in reality, etc.

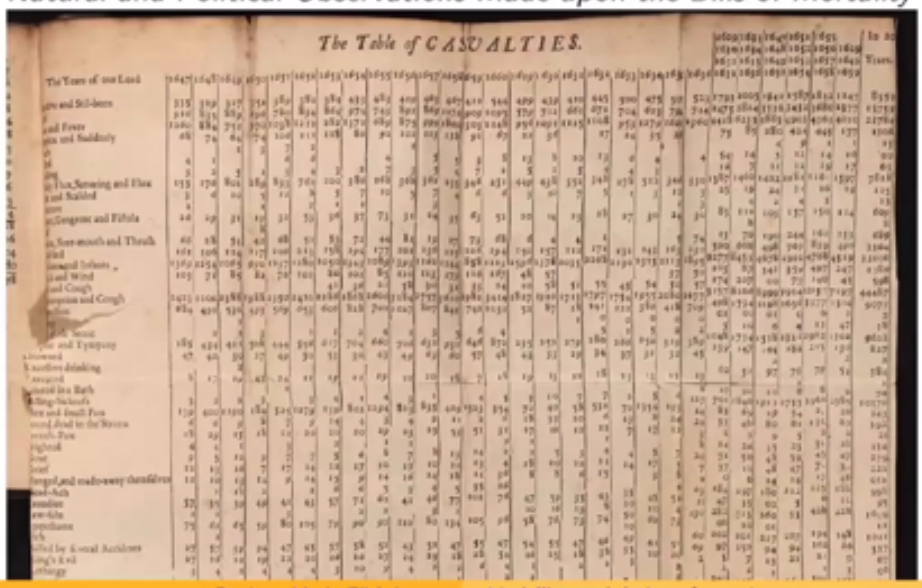
notes

summary

6m 2s

Graunt's book 1663

Natural and Political Observations Made upon the Bills of Mortality


The Table of CASUALTIES.

Metadata are very important. So the relationship between them is key. What does it really tell me in this case? So understanding that type and choosing appropriate statistical methods or ensure accurate analysis, importance of metadata. I already mentioned it and relationship. Okay, it was everything I tell you, it's written.

notes

summary

8m 31s

1.3.5 Categorization : Titanic passengers

		Dead		Survivors	
Age	Class	M	F	M	F
Children	1st	0	0	5	1
	2nd	0	0	11	13
	3rd	35	17	13	14
	Crew	0	0	0	0
Adults	1st	118	4	57	140
	2nd	154	13	14	80
	3rd	387	89	75	76
	Crew	670	3	192	20

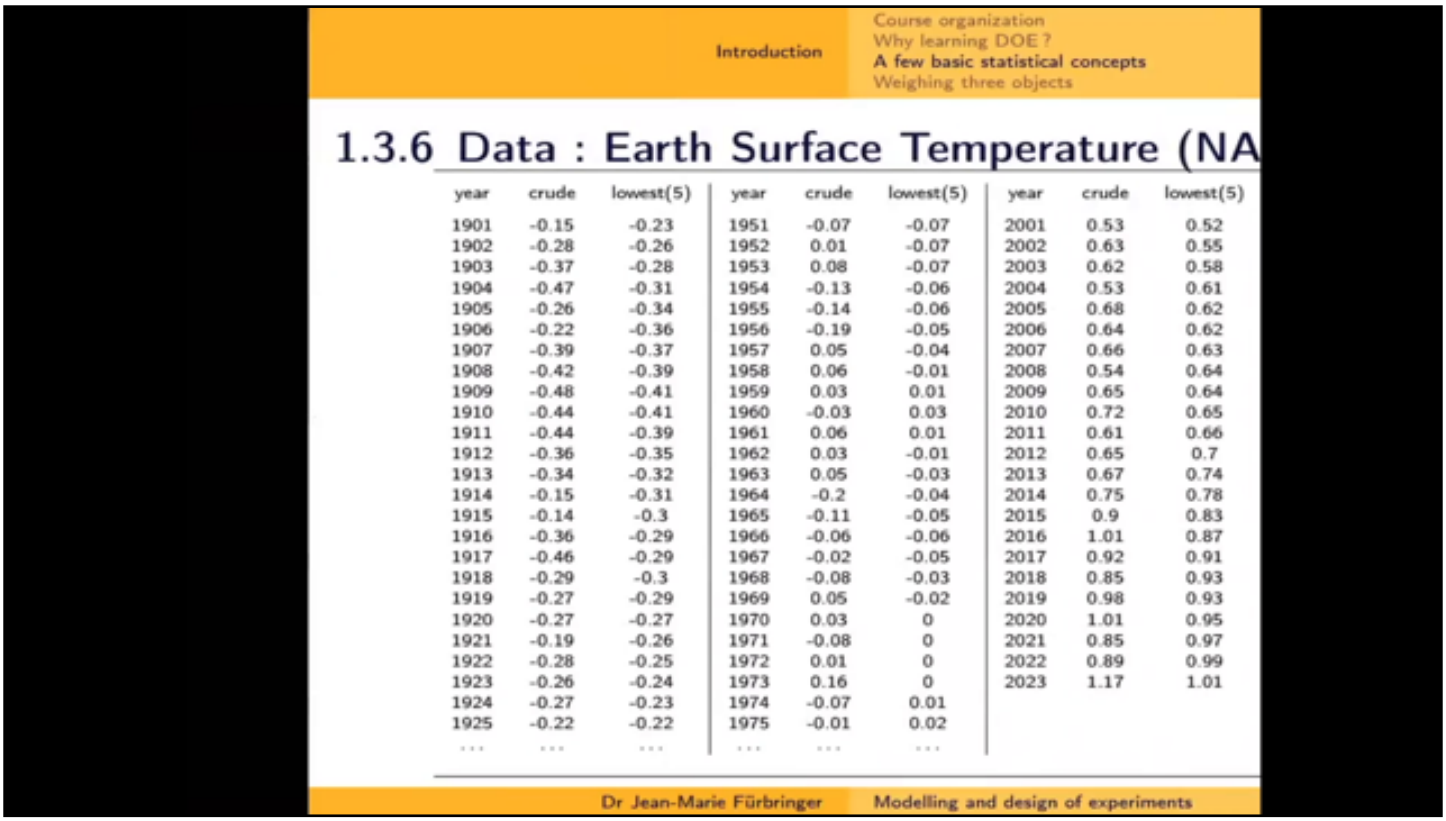
"[...] Those methods deal comprehensively with entire species, and with entire groups of influences, just as if they were single entities, and express the relations between them in an equally compendious manner. They commence by marshalling the values in order of magnitude from the smallest up to the largest, thereby converting a mob into an orderly array, which like a regiment thenceforth becomes a tactical unit. ", F. Galton, Biometrika, Volume 1, Issue 1, October 1901

So data, this is an example of very, very old data, perhaps first data for making statistics. It's about the bill of mortality in London or in England. Look at the years, 1640 something, etc. trying to understand. It's data that is not to prove that is also data that have been organized. And so we start to eventually understand something because we see frequencies being higher than others and eventually I don't know. Yeah, I cannot read that but those are people that died from gangrene. So you see the problem of gangrene is a very important health problem, etc. Data organized by categories is the first part of the analysis. Here is an example of the people who died in the Titanic. And observe that when you organize your data, you start to understand something that a story behind the organization of the data. Yeah, we have young boys and young girls have died also.

notes

summary

9m 1s



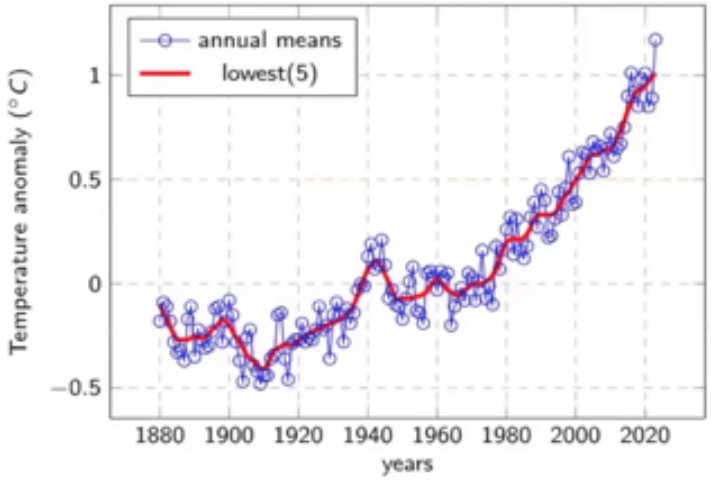
And after you can see who died or survive the crew, mainly the crew pay a very large bill to this accident. So organizing the data is a way of starting to understand things. Okay, I have put a text from Galton in biometrica. as in fact the statistic in your geometry. Those methods deals comprehensively with entire species and entire groups of influence, just as if they were single entities and express the relation between them in an equally compendious manner. They come as by more shallings, a value in order of magnitude of the organization of the data, from the smallest up to the largest, thereby converting a mob into an orderly array, which like, I like this sentence, which like a regiment, then forced to become a tactical unit. Written in 1901.

notes

summary

11m 13s

1.3.7 A picture is worth a thousand words

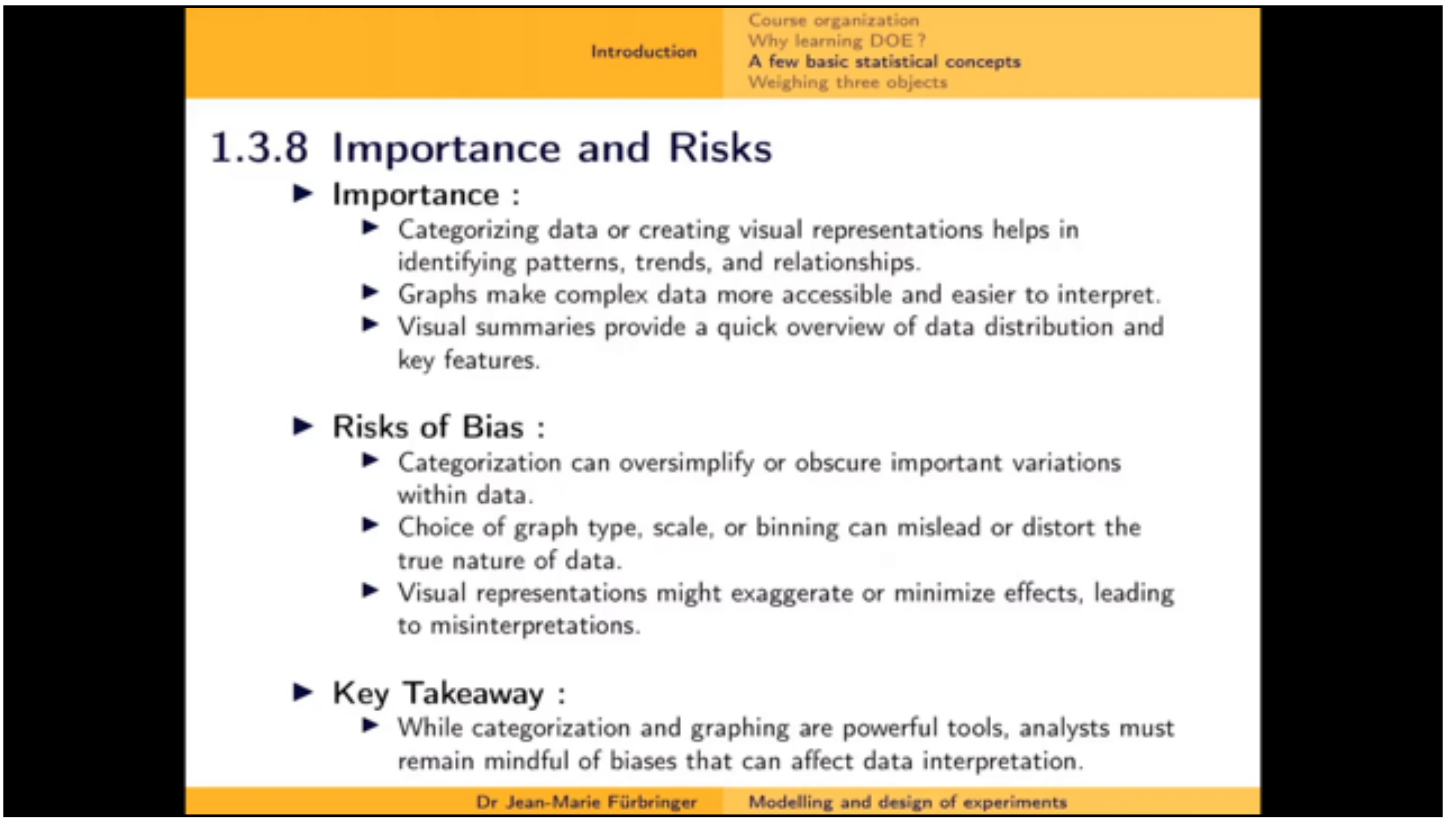


So here is another type of data. Something funny about my relational strategy, I had my data when one day or during the weekend, one day it's like, what could be the data I can present to my students for understanding what is important of data? It gave me a few and it gives me temperature of the earth, of the surface of the earth, the average temperature. And okay, you see the data and you are, it's not possible to classify, we classify by the date, it's a rolling date, but with the data, what is also very important

notes

summary

12m 52s



is to make graph. So okay, data classification, but also to make graphic is the way of getting inside of what is in your data. So you see in red, the, now perhaps in blue is the real average measurement, I didn't check, I think picture is worth, is worth a thousand words. It's very important to make graphic of your data and make the analysis of your data but also the graphics let you discover when you are very, what's happened is a trend that you can also see. provide a quick overview of data distribution and key features. Nevertheless, there are a risk of bias. Categorization can over simplify and obscure important variation within the data. Choice of rough type scale. You know that we can play with the scale, where is your region? Probably you want to talk to you in this discussion. Can mislead and distort the true nature of the data. Visual representation might exaggerate and minimize effects leading and misinterpretation. So make that categorization.

notes

summary

14m 8s

1.3.9 Pareidolia

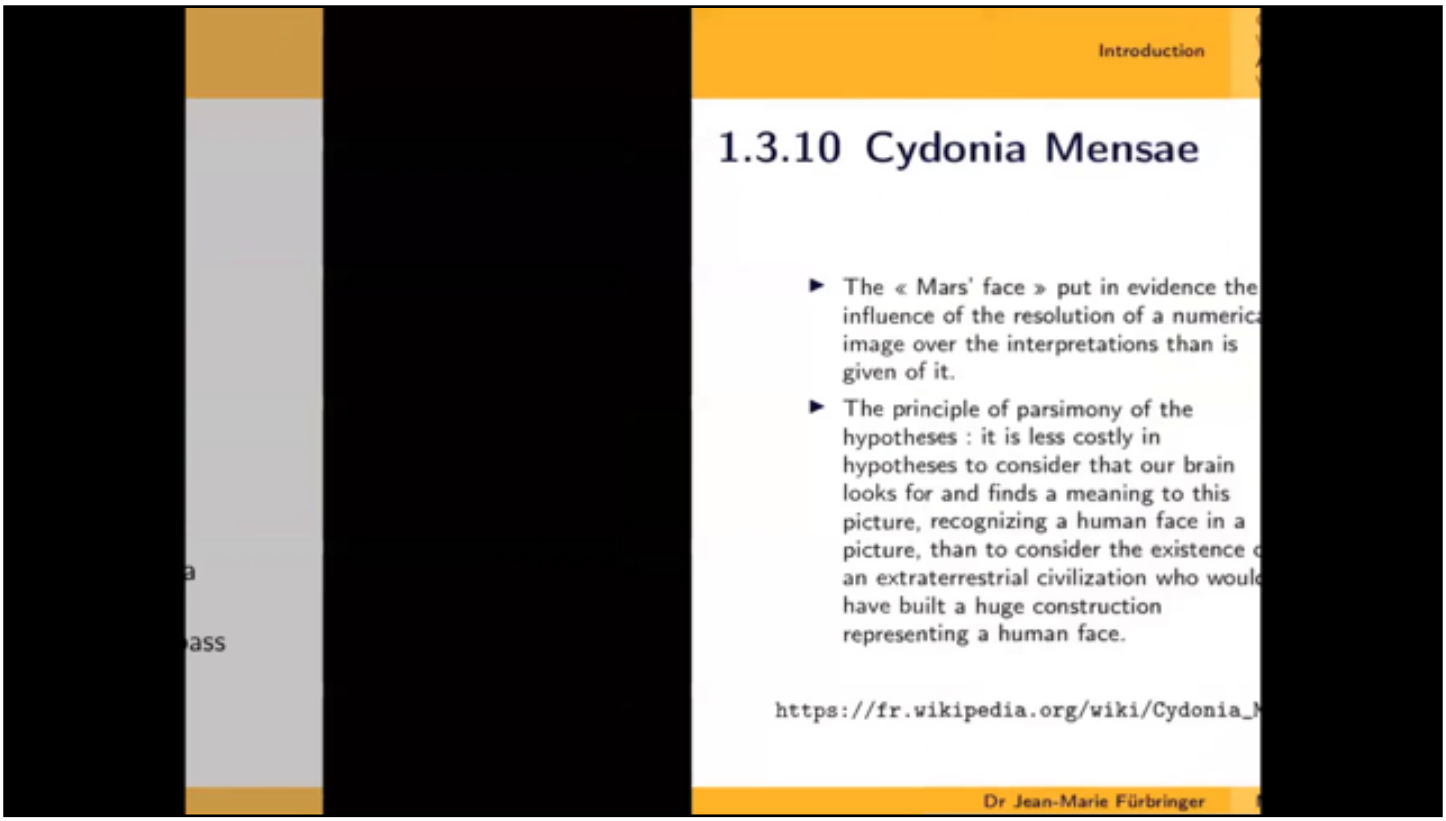
Our brain has the tendency to see motives in random sets as well as a tendency to make a script, which means to tell a story, and give a meaning to an image. Statistics offers tools, like statistical tests, to pass the perception.

Make graphic, but don't fall in the trap. You will see that the only thing that you have to see. So take away while categorization and graphing are powerful tool. Analyst must remain mindful of bias that can affect data in reputation. The story of this slide. So I have this in mind. I write it and I ask Chagipiti, what do you think about it? I'm really curious.

notes

summary

16m 49s



So I wanted to illustrate a few of the bias. If you know this, it's my favorite bias. It's called pare idolia. Idolia means an image. It is the fact that as human, we all the time interpret an image and we interpret centered on our existence. So everybody will see characters in this picture. My favorite is the washing machine. And this one about the electrical light. I was never, I never see a really light person. United States, this type, I never react, but this is very good. All are very good. It's also what makes the comic so powerful that you make a circle with two dots and everybody see a face.

notes

summary

17m 27s

1.3.11 Survivorship bias

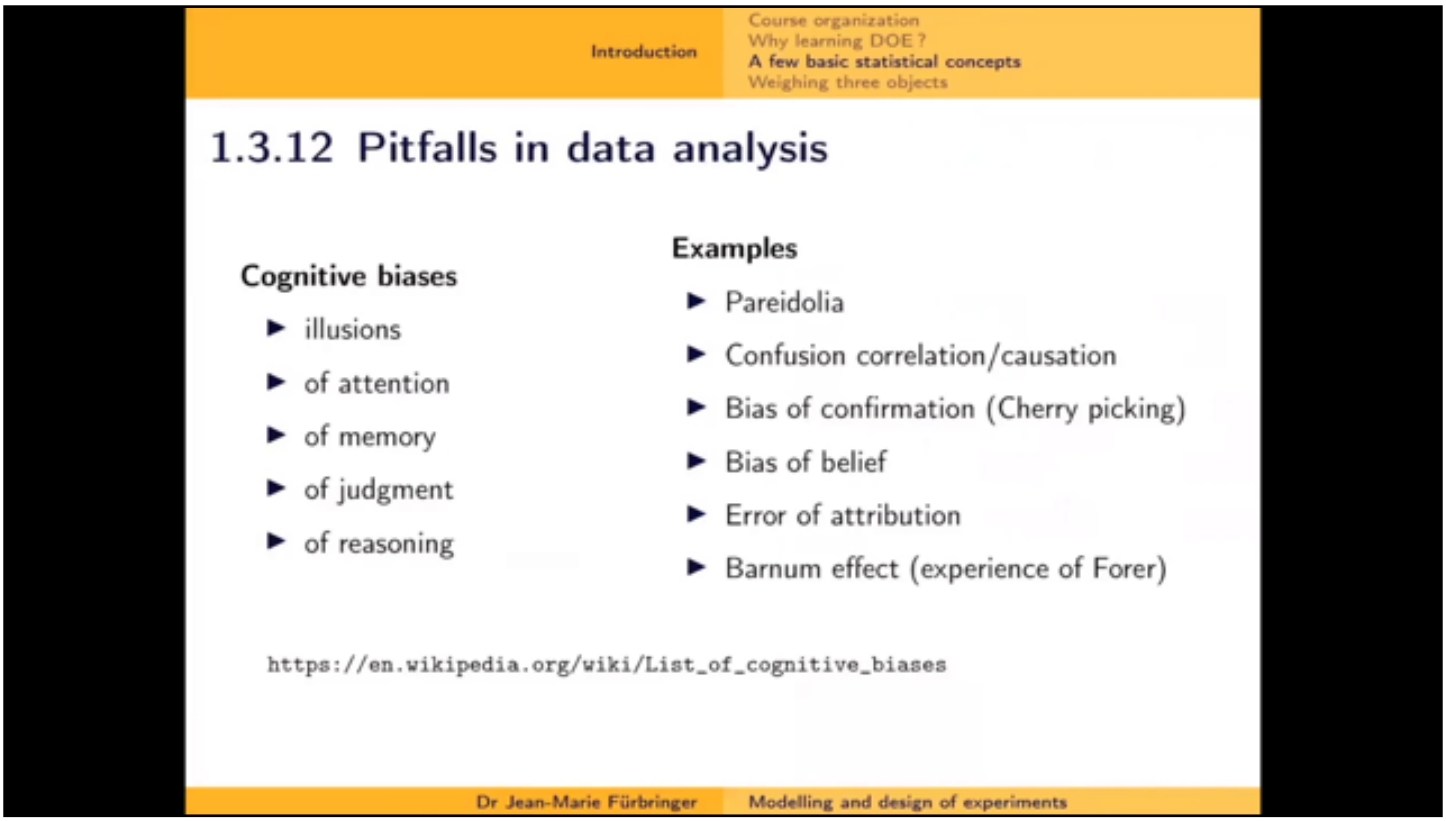
The damaged portions of returning planes show locations where they can sustain damage and still return home ; those hit in other places presumably do not survive. (Image shows hypothetical data.) The error was to consider that the planes are *more probably* hit in those places and then reinforce them.

It was just a question of the resolution of the picture. A few years later, when we were able to make better picture of Mars, the statue has disappeared. Opax immersion has erased when they see that the human way. They think that they are amazing. So this type of bias exists also inside. It's not only for making fabric and another reform for laughing, but it exists also in the reality. We have the impression of a curve say something or show us something. And it's better to check better and to see the resolution and things like that.

notes

summary

18m 48s

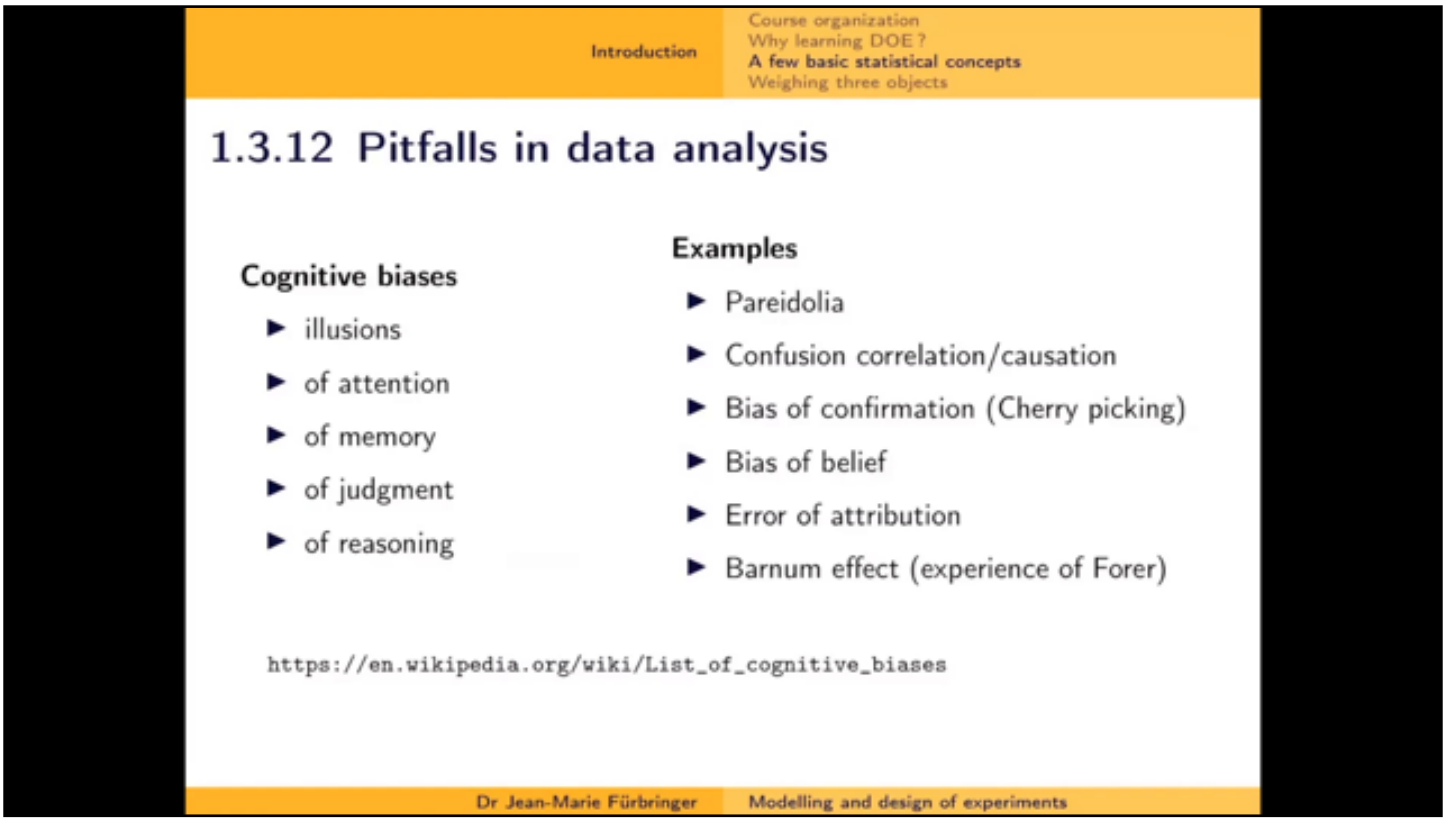


Another bias is all the survivor bias. In the second World War, British aircraft makers looked at the aircraft they come back from the battlefield and tried to see where they get peace. And the problem that they tried to make a better protection of those zones. But the product was about the thing that didn't come back. And it was not necessary to make those regions stronger. It's a region where no planes come back. I mean, here is a place. So OK, this is what we call the survivor bias. I remember when I made the interview for being the main deputy director of mechanical engineering. So the question was about women. They wanted to have more women in mechanical engineering. And I thought, what is the case today? But at my time, they are think or treat like in physics also for girls. And so in the interview, they asked me what you would do for understanding what is the program or you do not have more women in mechanical engineering. And I say we have to ask it's a girl who have got costs. Because if you ask the question, you're the guy here, we're the guy here. You don't have the answer why other and not come. Survivor bias.

notes

summary

19m 48s

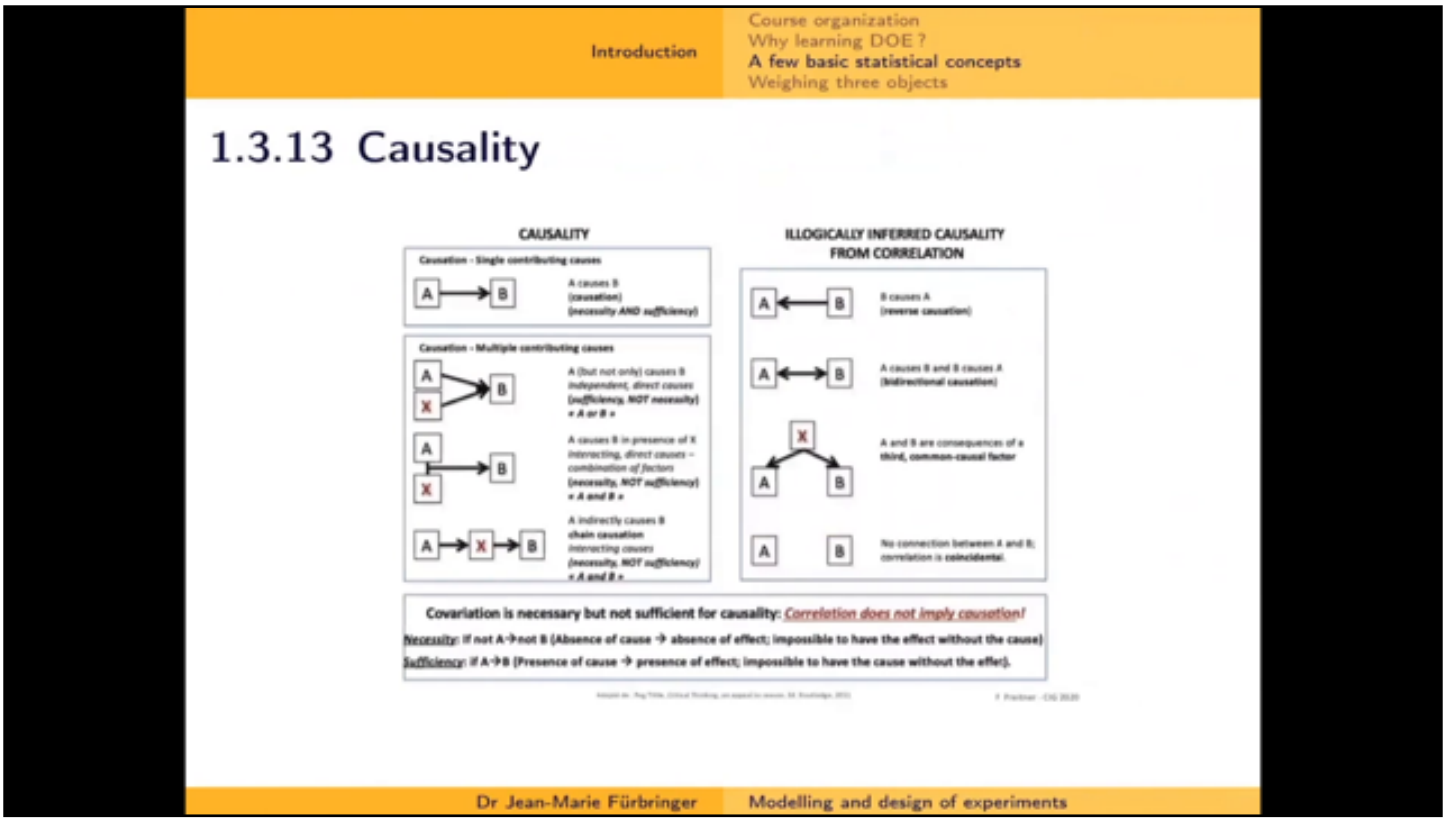


OK, there are a lot of biases. As a researcher, you have to be aware of these type of bias. So they can happen with your analysis. So we have illusions. We have bias of attention that we could more attention and something we know than something that we don't know. Of memory. Yes, I remember that. But a lot of students don't remember that it can be as important. Of judgment. And the specialist of judgment. This is important. This is not important. Even specialists and experts make errors and we can get elements of reasoning. All the problem of causation and correlations. And we mix of judgment and correlation.

notes

summary

21m 46s

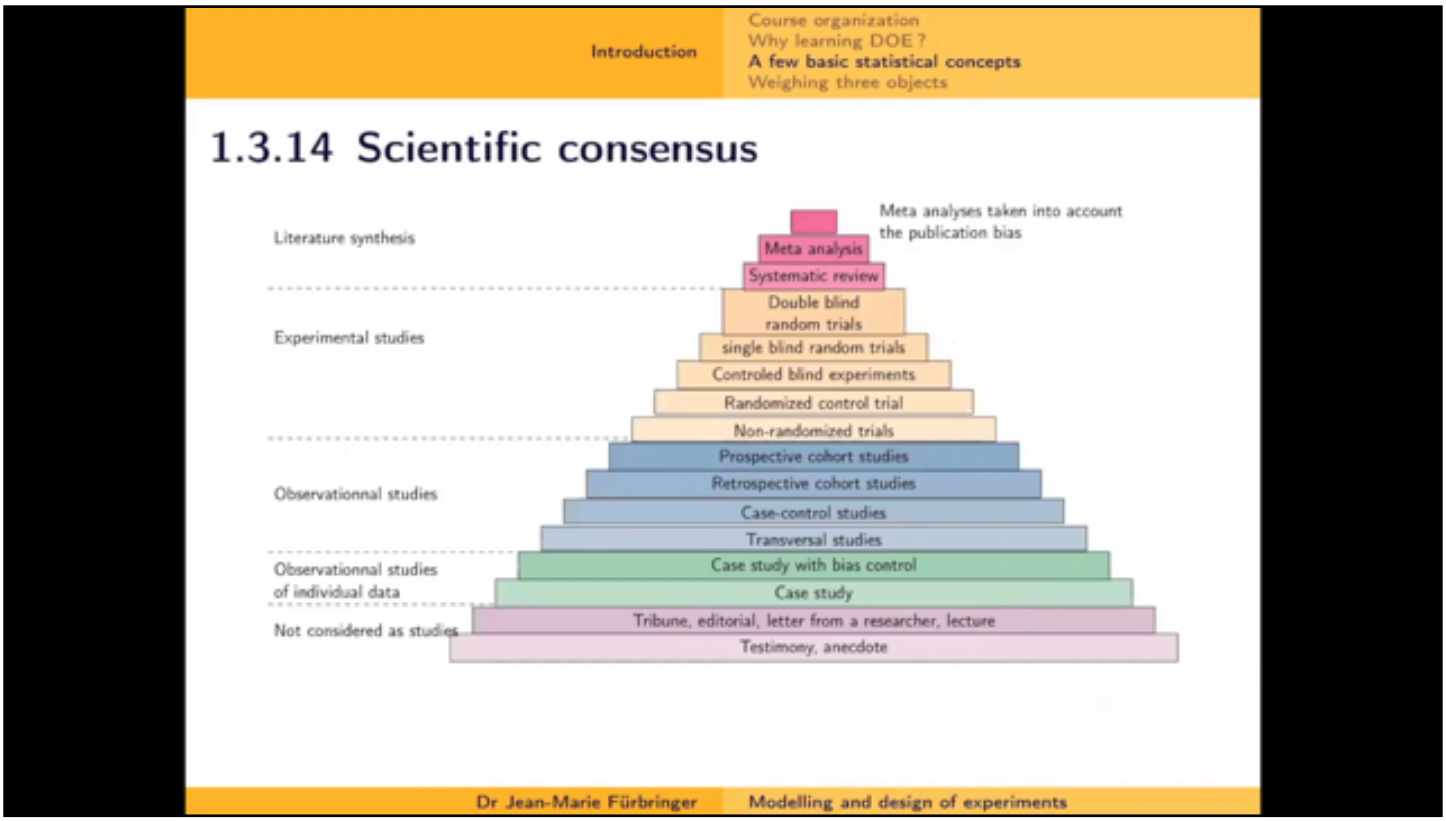


So here I put a few examples. Parahidolia, confusion, correlation, causation is a very classical bias. It's not because you see a current evolving at the same time. It's your effect that's really there. I think this is a cause. It's really these factors that you are checking. Perhaps both are evolving. I don't know if you ever see this correlation between the numbers of movie or Mr. Cage and the number of deaths by hanging in the United States. You see like that. So it's really possible to put correlation absolutely stupid. But why? Because they depend about the same factor. And so it had been named Barnum effect, because Barnum was a circus in the United States. If you are interested by that, a long list of committees are viable. A lot are quite closed, and sometimes the name doesn't make a big difference.

notes

summary

22m 37s

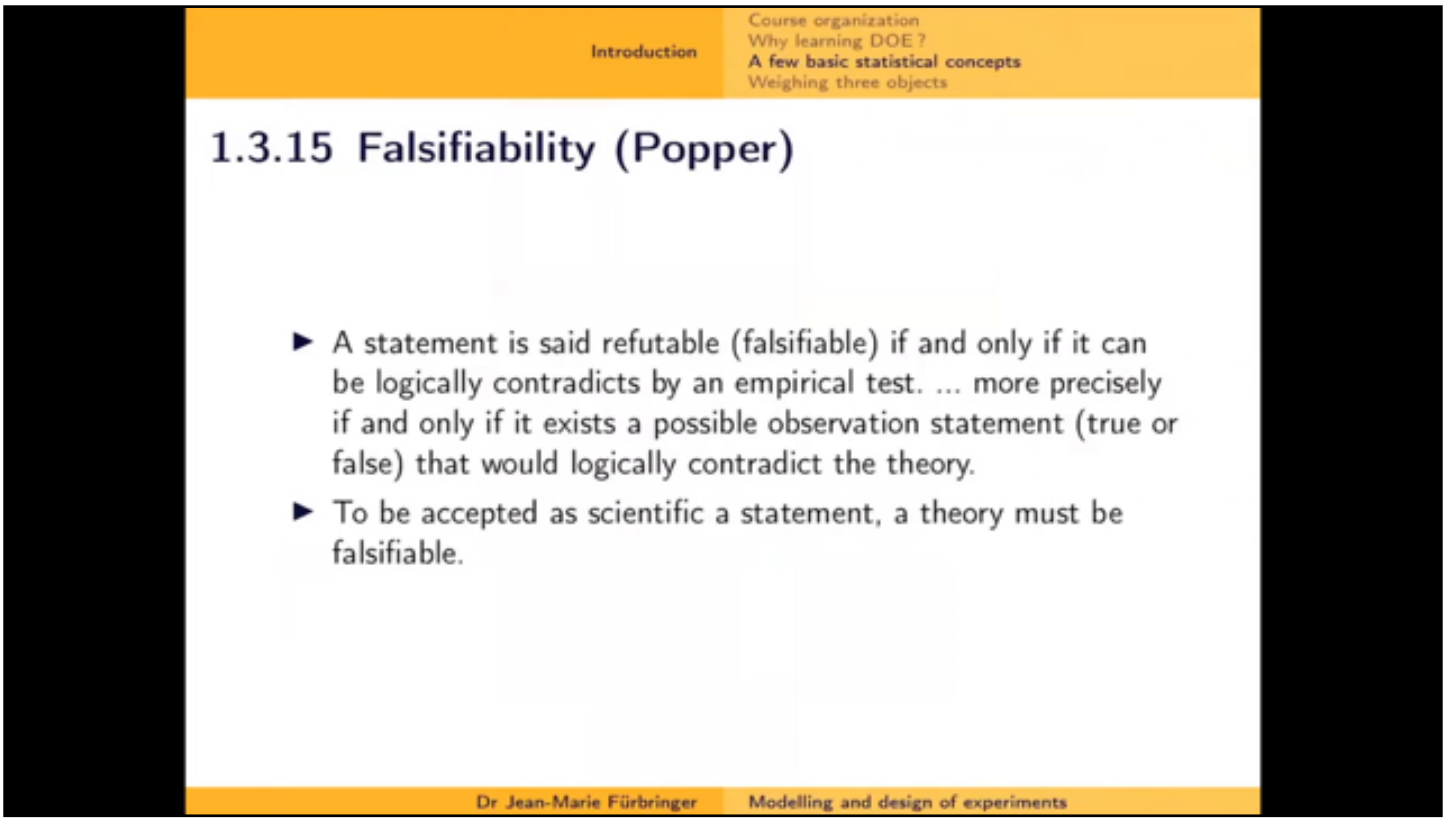


A is provoking another event B. And it can happen different things. It can happen that several events can also provoke the same thing. So be careful when you go to the US wide, because you can have different causes of something. It's not because you have proved something from A to B that all B to A will be the same. You can have several factors that have to be in coherence for provoking something. Or you can have some intermediary paintings or what's the X. After the bad thing that you can admit between cause and error, typically when there is feedback loop, it will be quite complicated to identify what is the cause and what is the consequence. You can have big directional causation when you have really an important loop. And you can have two events that seems to be related because they are depending on another. What I explained about the movies of Mr. Cage and the Death by Hanging. And sometimes things are, they look correlated, but in fact, they have no relation. It's wrong. And it could happen that something seems to be related. But in fact, there are no way of rating them. And finally, they are perhaps not related.

notes

summary

24m 49s

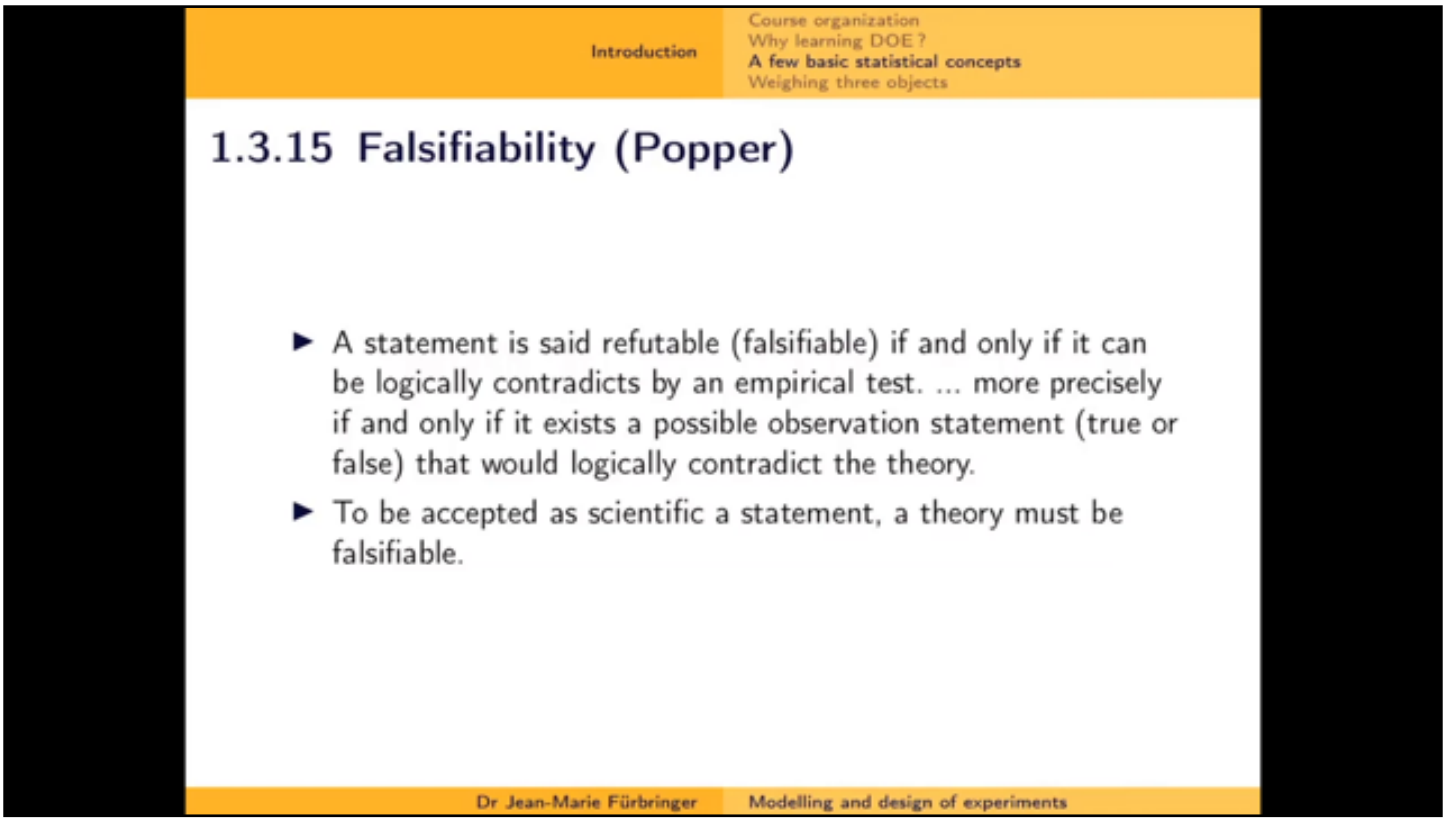


Causality. I want to be quick. And this is just a stairs about what we call the scientific consensus. Usually in science, like engineering and physics, we have not too much problem with that. In life science, in medicine, it became quite complicated. I will not talk so much, but you can understand that between the testimony of somebody, they say, oh, I see that. And having a meta analysis of different papers and all the literature in the domain, you have level of consensus. And when somebody threatens something, it's really true. It's quite interesting to understand that which level of the consensus. We are.

notes

summary

26m 44s

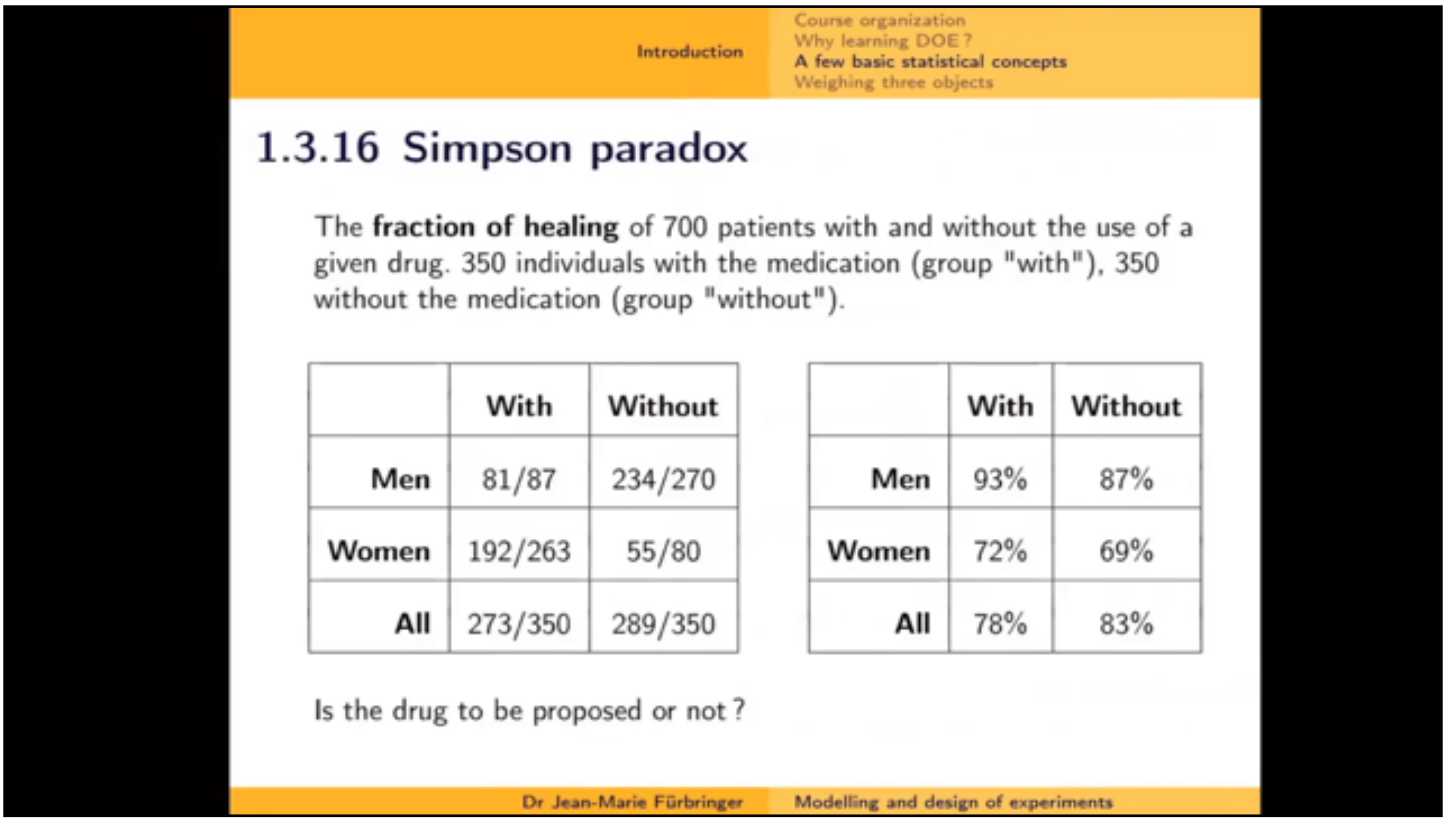


Poor being, science, a proposition needs to be passively able. The passively ability. That means that you need to be able to make an experiment that will accept or refute the proposition. If you are not able to do that, in fact, it's not science. So cosmology, perhaps, is not a science. Because there are limits of things.

notes

summary

27m 38s

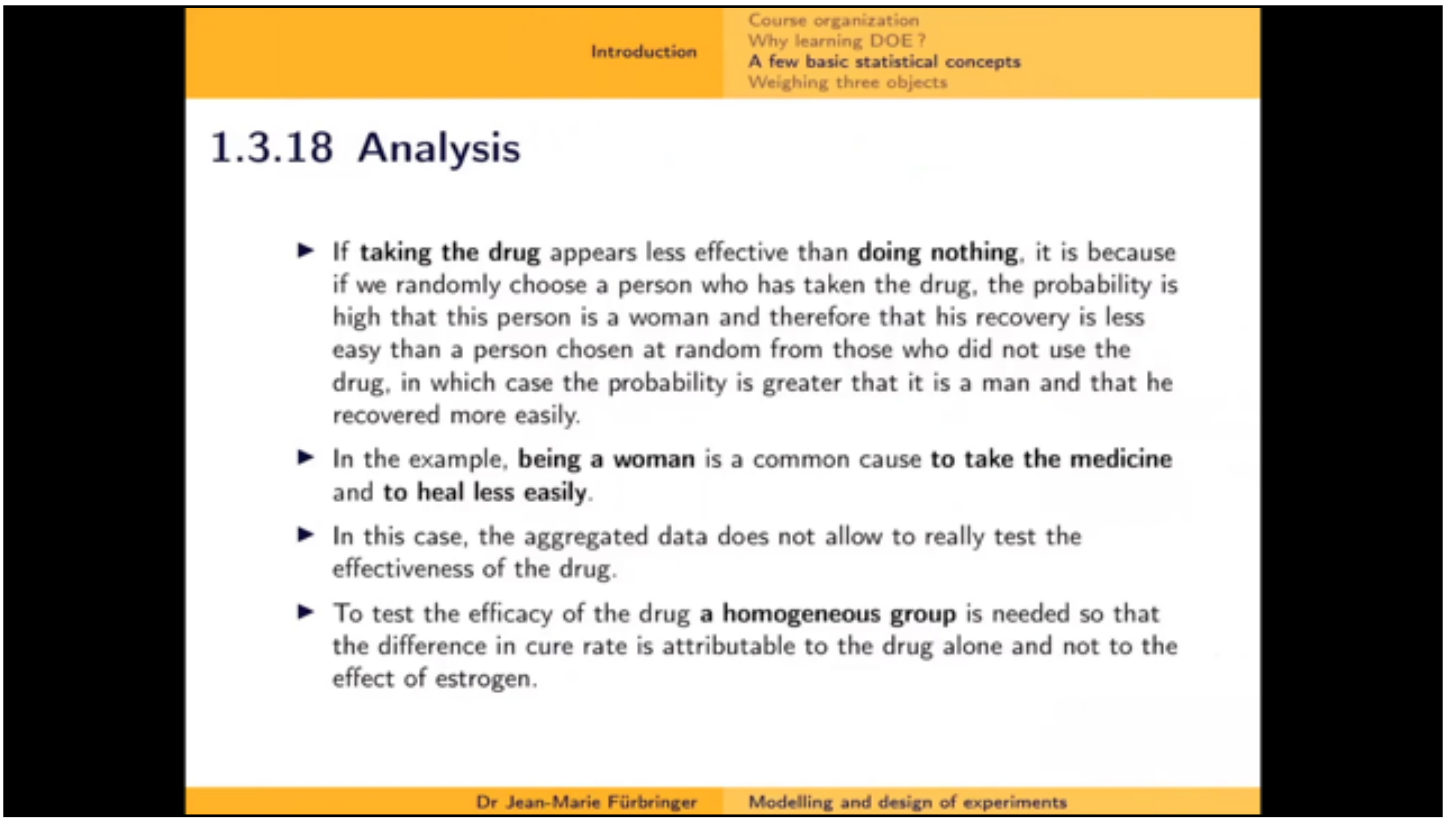


And no problem of accepting cosmology as a science, but you understand. Some people say, typically, psychology is not really a science. Because all psychology, all the element of psychology, that's all. You cannot put people in an experiment to see if they became here. It's not accepted. So to be accepted as a scientific statement, a theory must be passively able. That means that you need to be able to make an experiment depending on the reason of the experiment between the two rules.

notes

summary

28m 13s

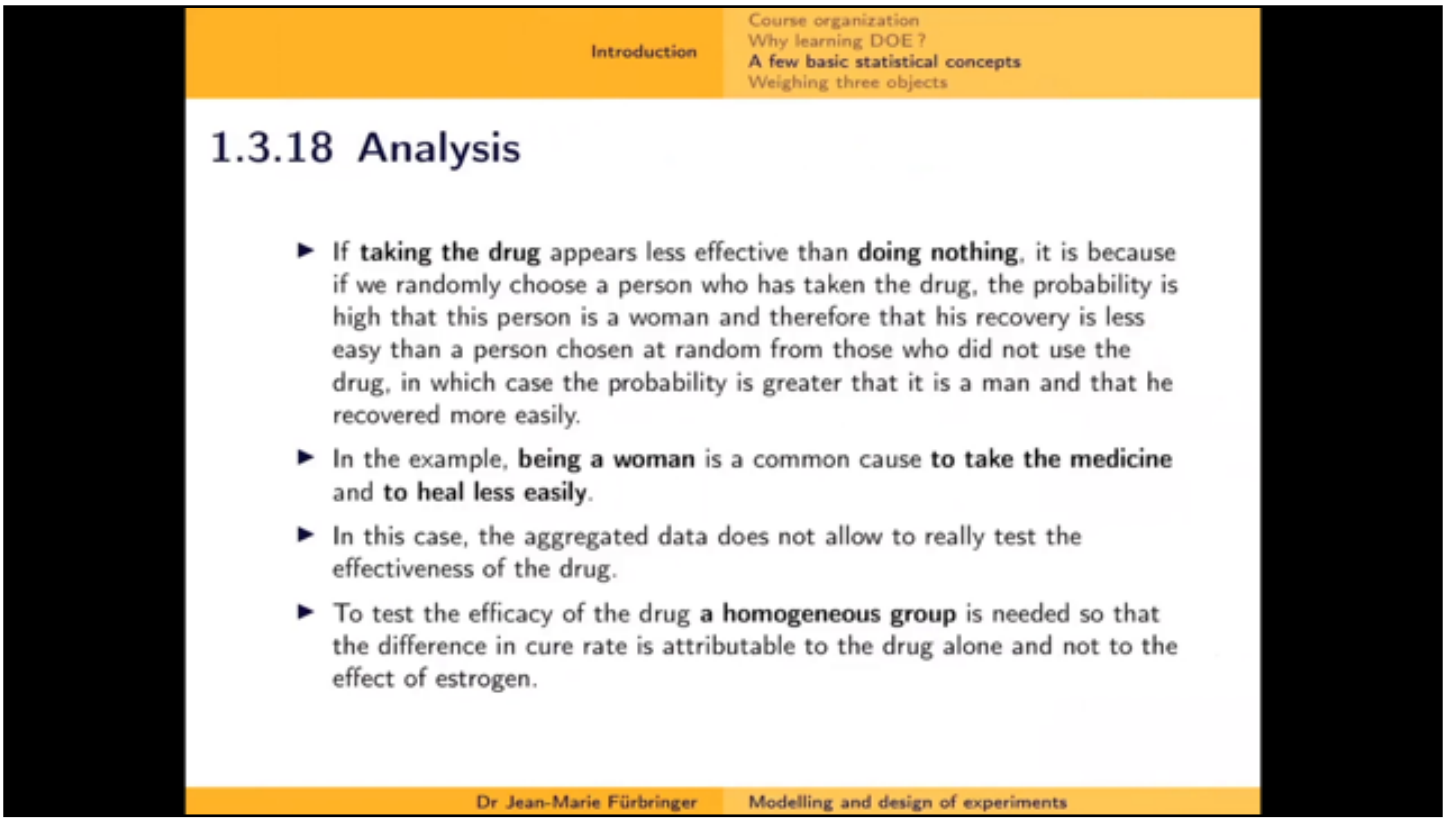


It comes from Popper. Simpson Paradox. It's also not a bias. A paradox is something that seems very, very particular. So you have 1700 patient, and you have a selection of 350 that have taken the drive. And 3002 have not taken the drive. So let's see, theory can be agreed with that. Let's see if the drug is efficient. see what is the ratio of creation for people who have taken the medicine in relation to people who have not taken the medicine. So on the 350 people who have taken the medicine, 273 that have taken the medicine have cured, have been cured. And when you look at the people who have not taken the medicine, you have even more, you have 289 that have been cured. So you say, finally the medicine is not really, when you look at that, you say, the medicine is not effective. And after you look at the man, and you see that in the man, if you look the one who have taken the medicine, in fact, you can look at the ratio if you want, you have 93% of the man who have taken the medicine that have been cured in comparison of only 87% of the man who have not taken the medicine that have been cured. And when you look at the woman, you see that you have 72% of the woman who have taken the medicine that have been cured, again only 69. So if you know the gender of your patient to give the medicine, but if you don't know the gender, it's a paradox, it's exactly a paradox. And the paradox is coming, if you check, because you have not an adequate numbers of person in each subgroups. And nevertheless, the question arises, do you have to put all the groups together or you analyze the groups in between? So if you want to solve

notes

summary

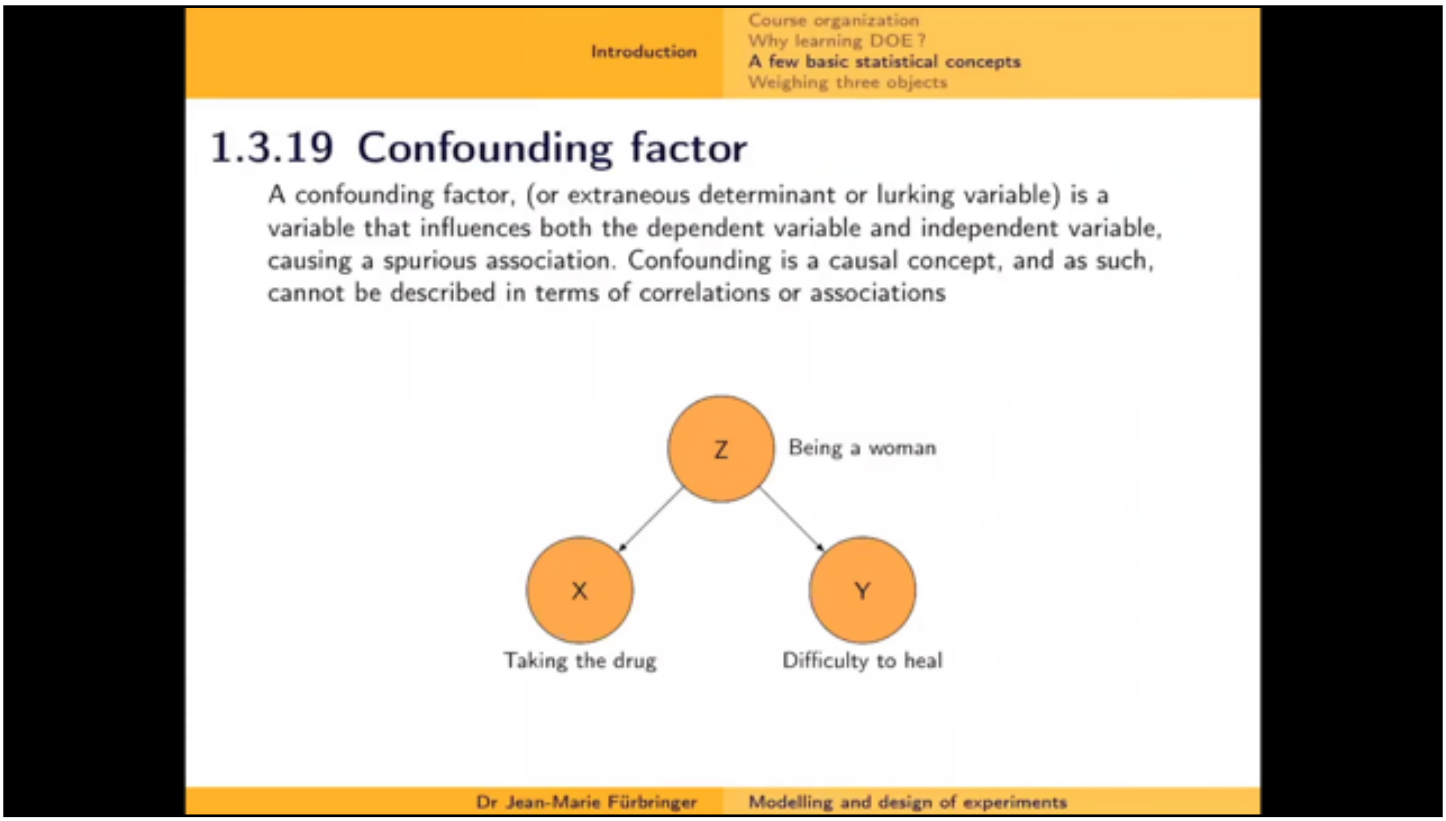
28m 50s



that, it's very important to make the calculations step by step. It's what I call a causal analysis, a cause analysis. How many people so we make for women and we compare them. We have an hypothesis, self-healing against medicine. So let's count really the same. It's what I'm doing is this type of practice. And when you do that, it's became clear that the situation, the story for making the long story short, in fact, the illness is related to this different movement for man, and it was not correct to make a group of all of that because it was different. But if you don't analyze carefully, you have the impression that you have a paradox. So in the exercise, you will have another situation, not between man and woman, but between pressure, the measurement of pressure, the same case that related to the measurement of pressure, how people, depending on the pressure, are cured or not by the

notes

summary



medicine. So here you have an analysis. If taking the drug appears less effective than doing nothing, it's because if we randomly choose a person who has taken the drug, the probability is high that this person is a woman and therefore that the recovery is less easy than a person shooting at random from those who did not use the drug, which in this case would be man. So this is the problem, the groups who are not of the same sex. In the example, being a woman is a common cause of taking the medicine and to heal. The two things, the probability to take the medicine was related to the probability to heal. In this case, the aggregated data does not allow to really test the effectiveness of the drug. To test the efficacy of the drug, an homogenous group is needed so that the difference in curing rate is attributable to the drug alone and not to the effect of being a woman. Now this is a case of checking the data correctly and be sure when you are making groups or not making groups what you are doing and why you are making groups or not making groups.

notes

summary

33m 49s

1.3.20 Data about scientific publications

	Country	Code	Region	Income	Level	2 000	2 018	pop2000	pop2018
2	Afghanistan	AFG	South Asia	Low income	1	4	112	20 779 957	37 171 922
3	Angola	AGO	Sub-Saharan Africa	Lower middle income	2	7	90	16 395 477	30 809 787
4	Albania	ALB	Europe & Central Asia	Upper middle income	3	22	180	3 089 027	2 866 376
5	Andorra	AND	Europe & Central Asia	High income	4	-	4	65 390	77 008
6	United Arab Emirates	ARE	Middle East & North Africa	High income	4	330	3 145	3 134 067	9 630 966
7	Argentina	ARG	Latin America & Caribbean	Upper middle income	3	4 386	8 811	36 870 796	44 494 502
8	Armenia	ARM	Europe & Central Asia	Upper middle income	3	346	521	3 069 587	2 951 741
9	Antigua and Barbuda	ATG	Latin America & Caribbean	High income	4	0	6	76 007	96 282
10	Australia	AUS	East Asia & Pacific	High income	4	23 276	53 610	19 153 000	24 982 688
11	Austria	AUT	Europe & Central Asia	High income	4	6 577	12 362	8 011 566	8 840 521
12	Azerbaijan	AZE	Europe & Central Asia	Upper middle income	3	155	761	8 048 600	9 939 771
13	Burundi	BDI	Sub-Saharan Africa	Low income	1	2	21	6 378 871	11 175 379
14	Belgium	BEL	Europe & Central Asia	High income	4	9 723	15 688	10 251 250	11 427 054
15	Benin	BEN	Sub-Saharan Africa	Lower middle income	2	43	228	6 865 946	11 485 035
16	Burkina Faso	BFA	Sub-Saharan Africa	Low income	1	48	252	11 607 951	19 751 466
17	Bangladesh	BGD	South Asia	Lower middle income	2	440	3 135	127 657 862	161 376 713
18	Bulgaria	BGR	Europe & Central Asia	Upper middle income	3	1 653	3 311	8 170 172	7 025 037
19	Bahrain	BHR	Middle East & North Africa	High income	4	83	322	664 610	1 569 440
20	Bahamas, The	BHS	Latin America & Caribbean	High income	4	2	20	298 045	385 635
21	Bosnia and Herzegovina	BIH	Europe & Central Asia	Upper middle income	3	75	704	3 751 176	3 323 929
22	Belarus	BLR	Europe & Central Asia	Upper middle income	3	1 170	1 180	9 979 610	9 483 499
23	Belize	BLZ	Latin America & Caribbean	Lower middle income	2	1	9	247 310	383 071
24	Bolivia	BOL	Latin America & Caribbean	Lower middle income	2	39	103	8 418 270	11 353 140
25	Brazil	BRA	Latin America & Caribbean	Upper middle income	3	12 783	60 148	174 790 339	209 469 320

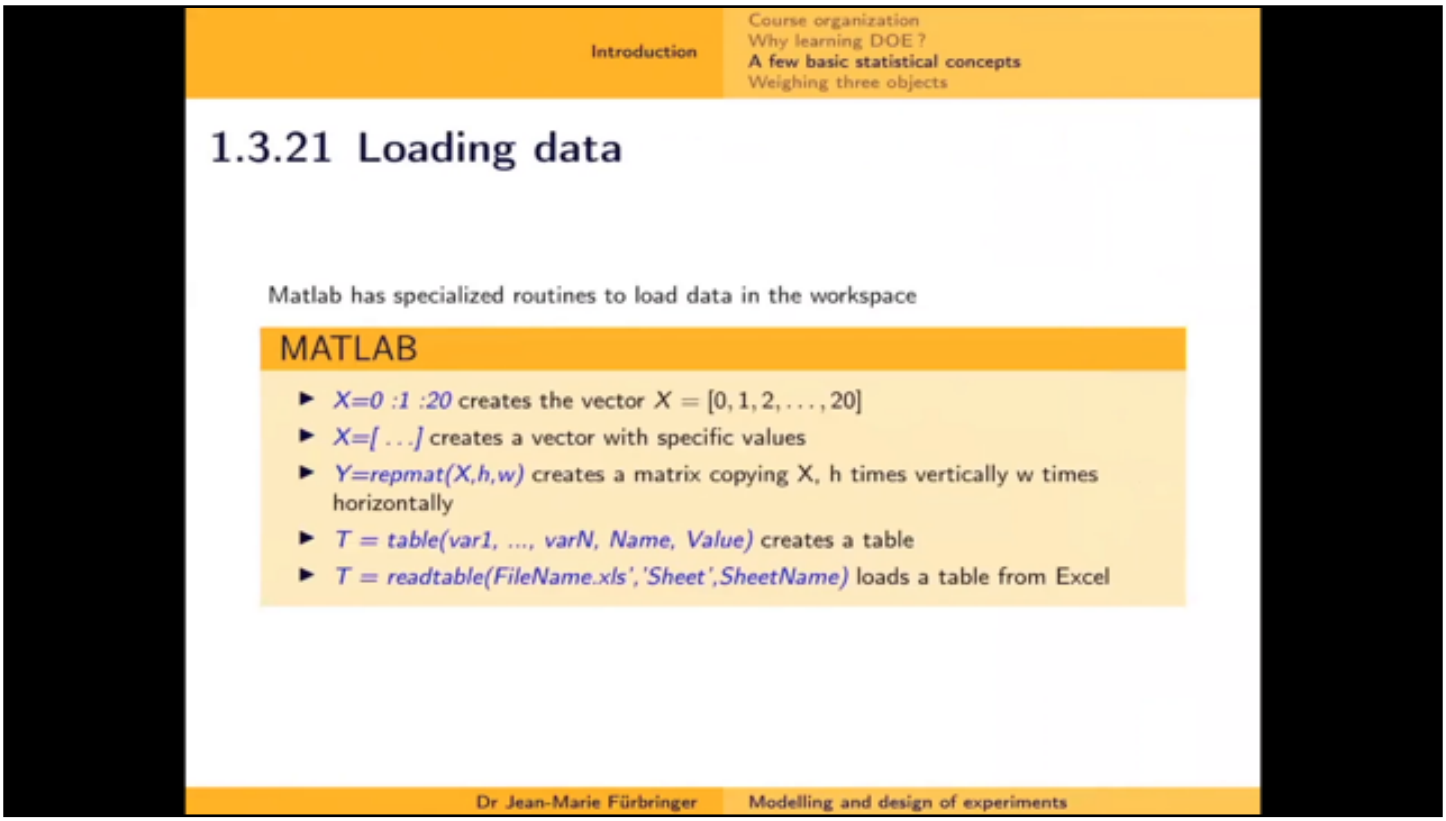
https://data.worldbank.org/indicator/IP.JRN.ARTC.SC?end=2018&start=2000

And there are all these ideas of a confounding factor. This is the concept in this situation

notes

summary

35m 6s



to be considered. So I'm proposing you a database about scientific publication and to check which countries are participating more in the science and technology.

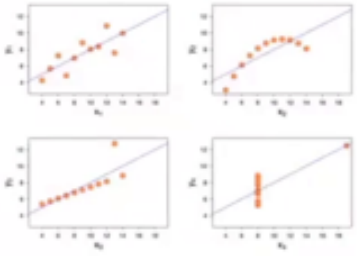
notes

summary

35m 20s

1.3.22 ANSCOMBE'S QUARTET

- Four data sets that have nearly identical simple descriptive statistics, yet have very different distributions and appear very different when graphed.
- Each dataset consists of eleven (x,y) points
- Constructed in 1973 by the statistician Francis Anscombe to demonstrate both the importance of graphing data when analyzing it
- Intended to counter the impression among statisticians that "numerical calculations are exact, but graphs are rough."



And I want to present you a few elements. So first thing is about loading data in MATLAB. You have some routine for loading data and you have here a summary, a cheat sheet for

notes

summary

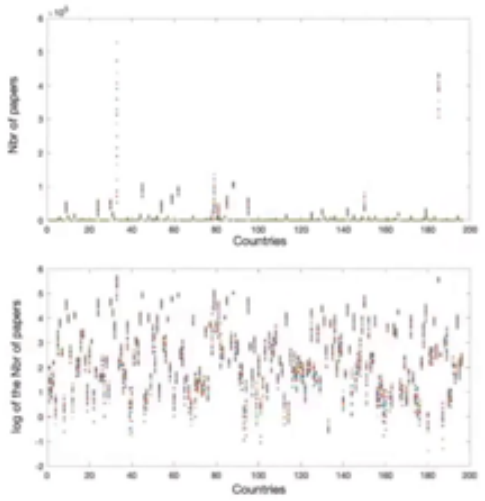
35m 33s

1.3.23 Visual analysis

- To detect patterns, aberrations, etc.
- The change of the metric is also interesting

Matlab

- *plot(x, y)*
  plot(x, y, LineSpec)
  plot(x_1, y_1, . . . , x_n, y_n)
  plot(. . .,Name,Value)
- *bar(x, y)*
  bar(. . ., Width)
  bar(. . ., Style)
  bar(. . ., Name, Value)

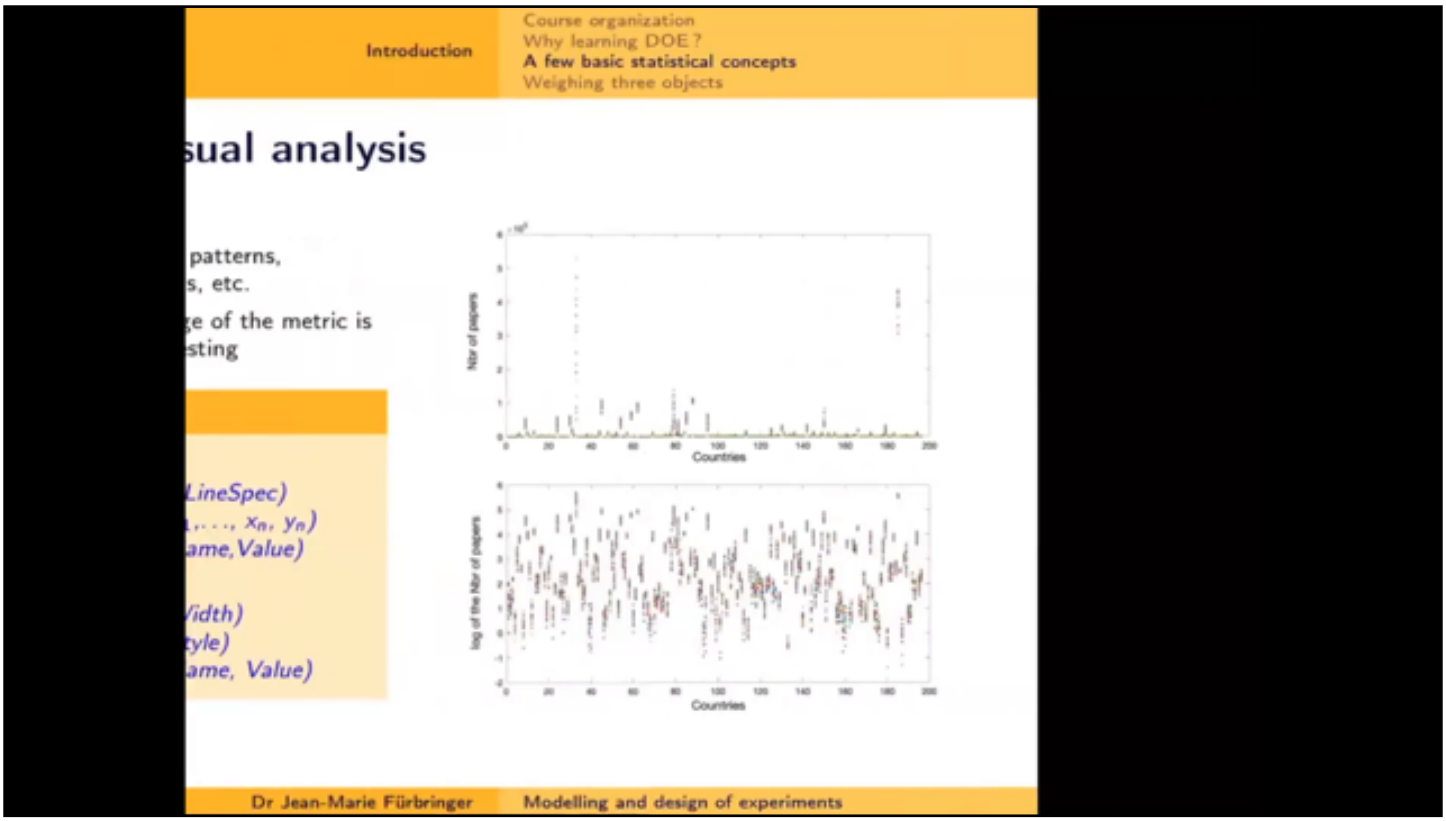


loading data in MATLAB. When you are making graphics, this is an argument also about buyers. I was hesitating where to put these slides. I would just let you know about something we call the Atcon Quartet. It's not a play of music. It's the fact that we need graphics and we need data and we need sense. So you see here four series of data. One is really for the way in which we are aligned. Another is in fact, so don't agree with relation. One has aligned with one point and is one point which is very different. And the other was a vertical line is one point very different. If you do the calculation only and never graphic them, they have the same risk, the same linear regression factors, the same quantity, you have the impressions, they are the same. This is what's just an argument for saying, make calculation, make graphics. That would not be the case. What is the first interpretation you've got when you see this is the real answer is you have a resonance at one point you have a resonance if I don't know what

notes

summary

35m 55s

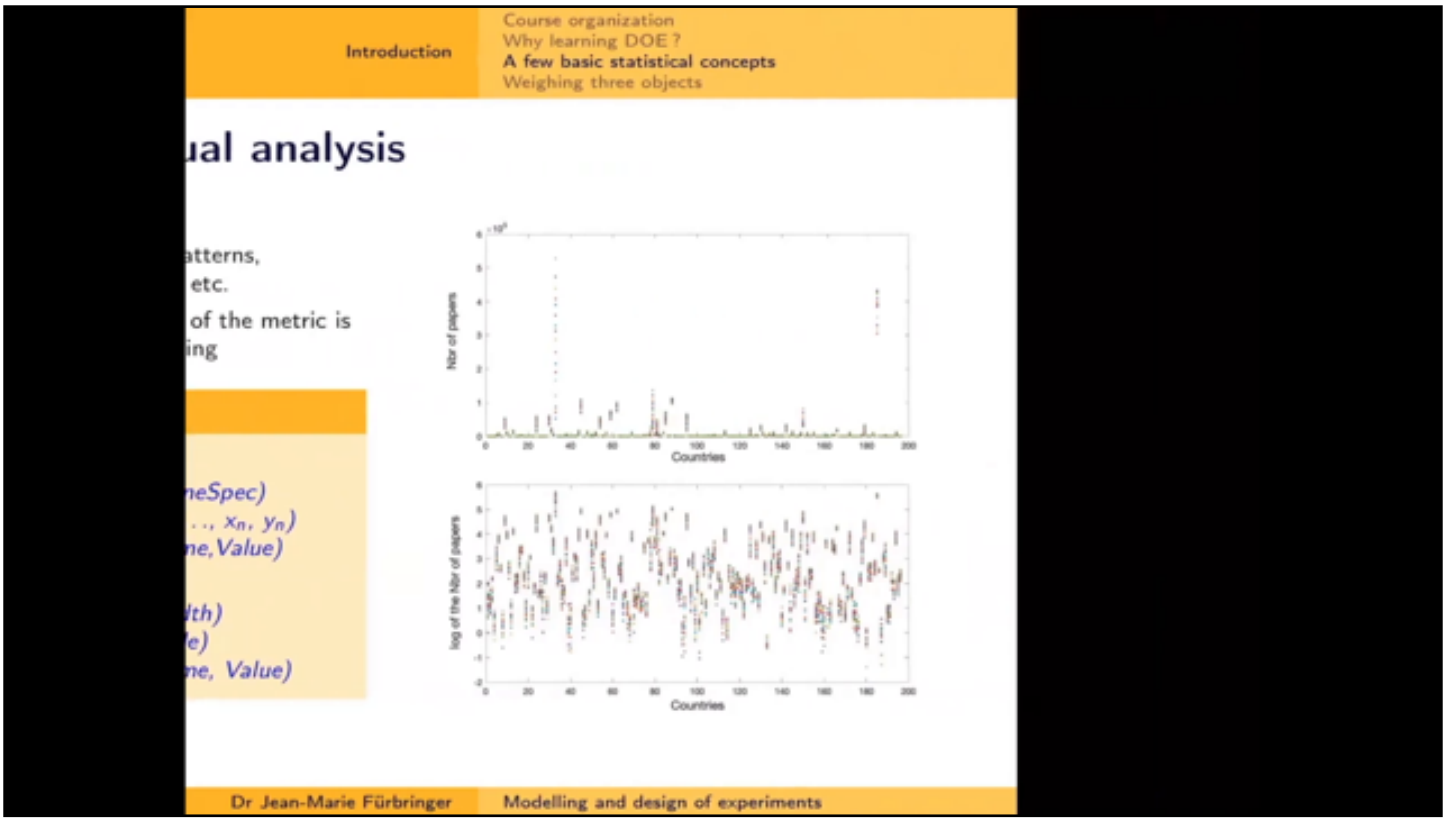


is it about it could be anything. For making visual interpretations or plots, which is one of the first functions that we check when we when we discover a tool for data analysis so you can make plots you can make bar with my plan if you are in python you can do the same very you learn things after it's important also to sorting the data because

notes

summary

37m 55s



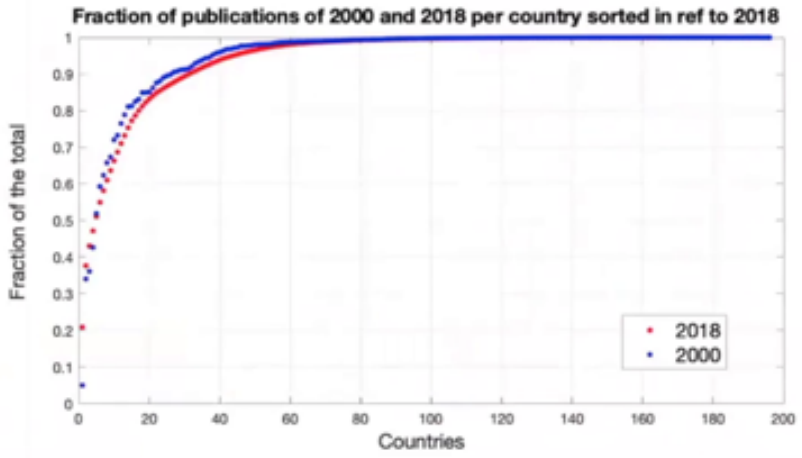
I look between this data about the publication okay you see to China and United States there

notes

summary

38m 25s

1.3.25 Plot of the sorted data

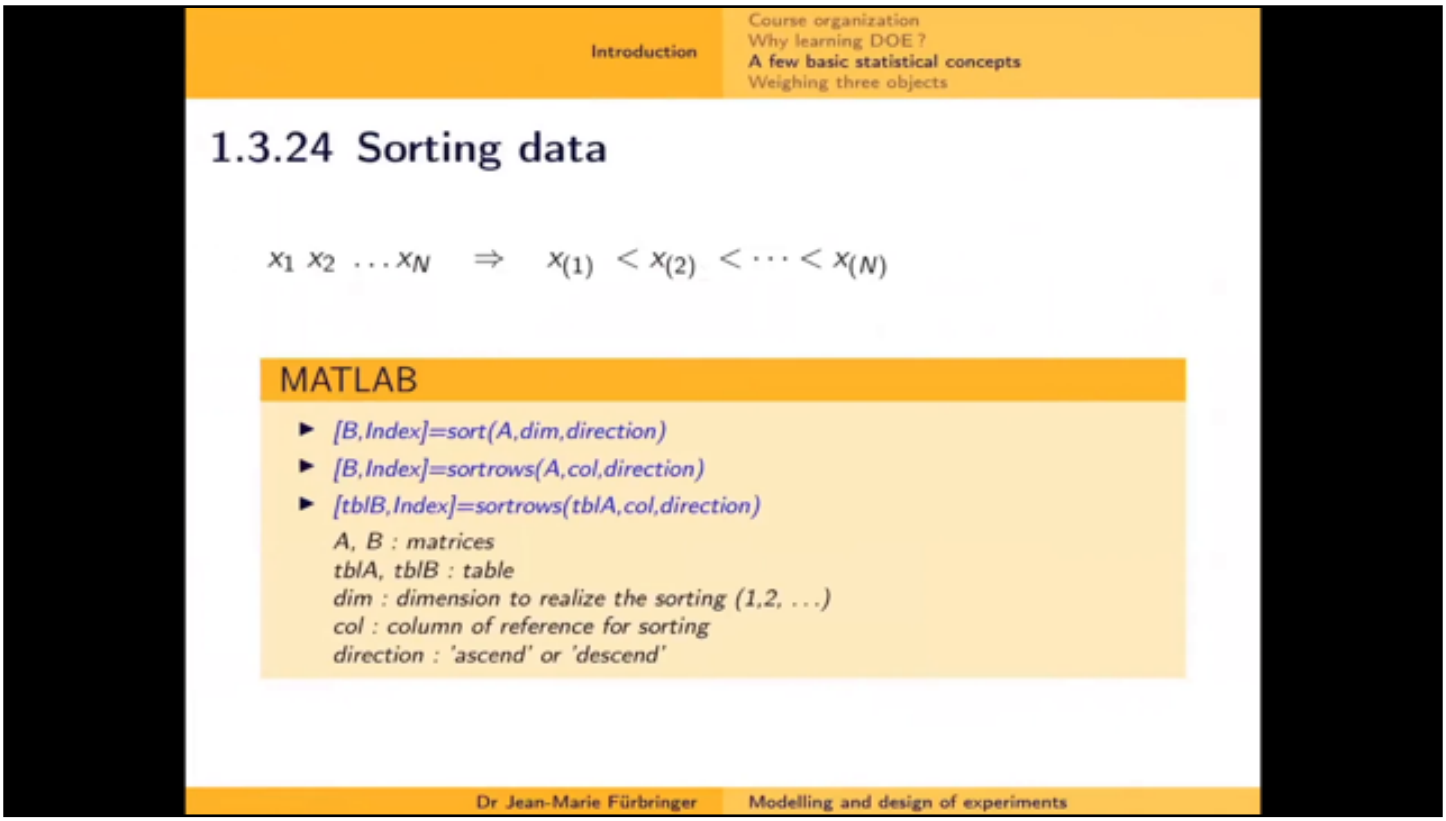


were really a lot of publication and but if you sort your data typically it was sorting by

notes

summary

38m 37s

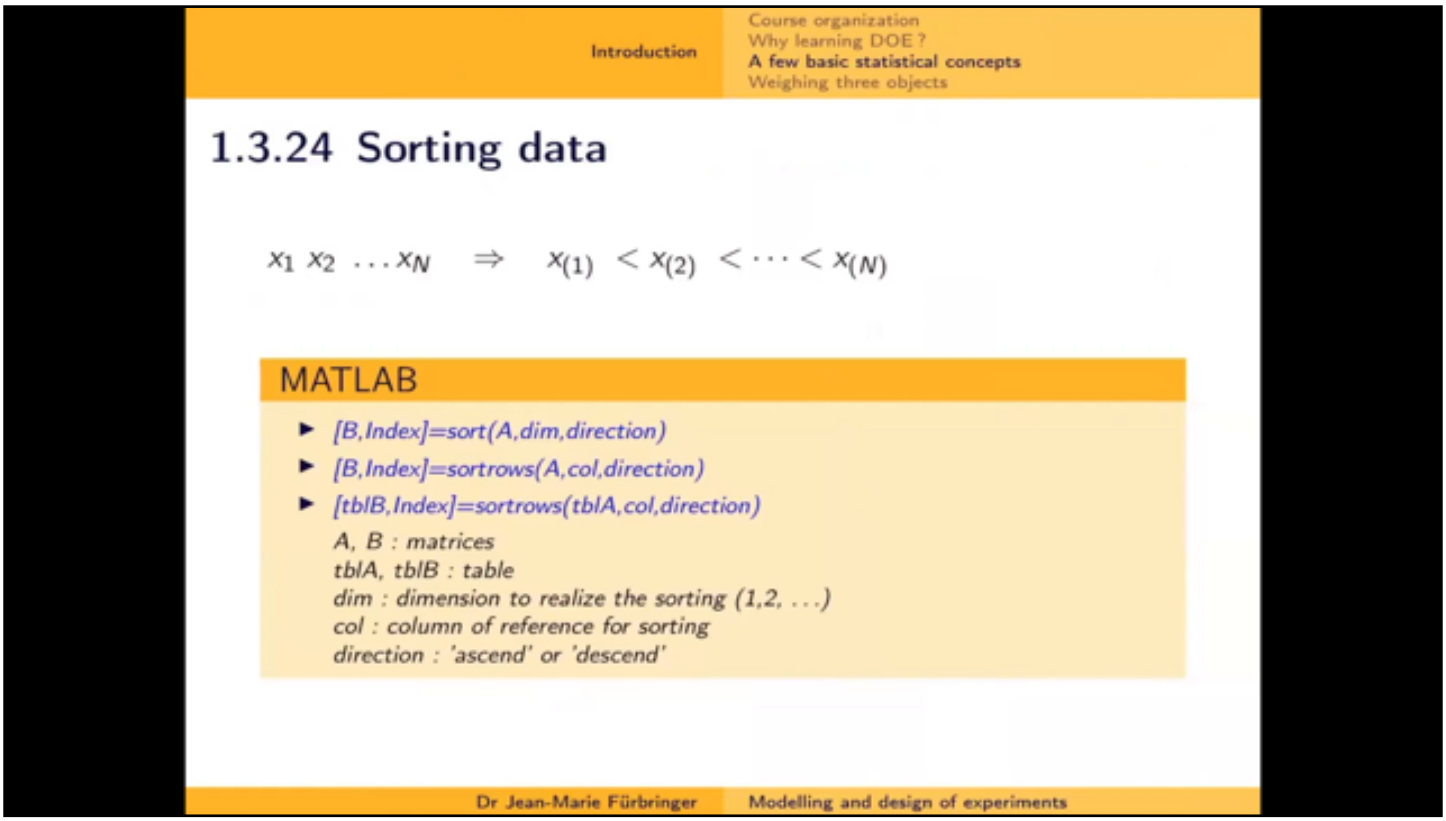


by what to do I sorting sorting the country from the from the highest one of the data we are checking from the highest value to the lowest value you start to see to see something that happened so you can tell us history here between 2000 and 2008 an evolution that you cannot see

notes

summary

38m 39s

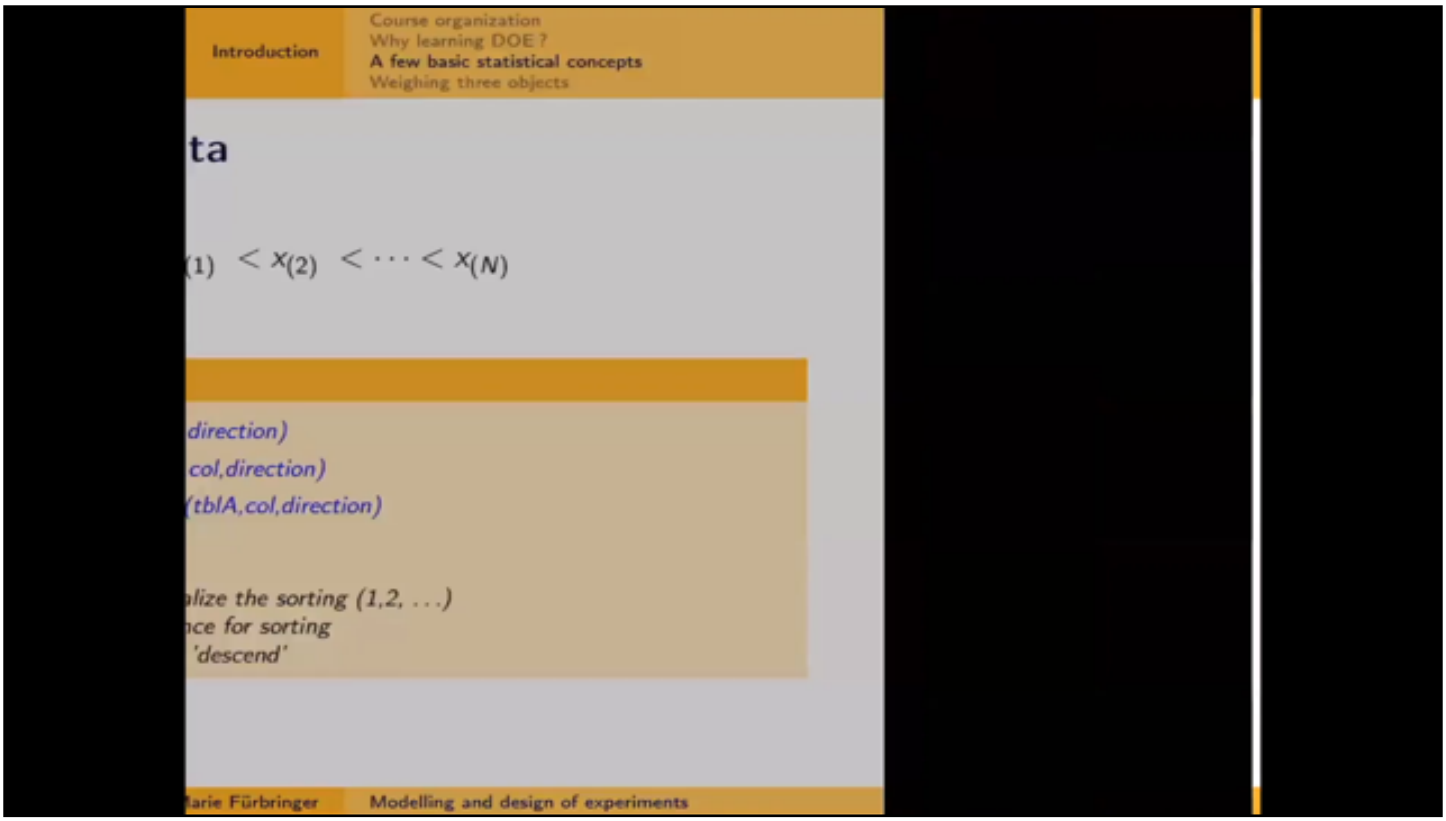


when you are putting your data without this sorting of the of the values where all the country are sorted I don't know by continent or by letter or by the first letter of the thing so sorting is quite very important but again it could be also a bias

notes

summary

39m 8s



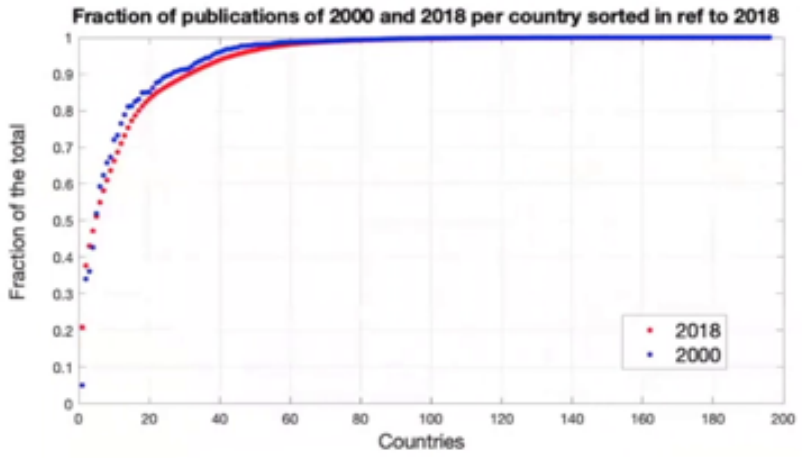
and here you see a few a few routines about the function for sorting so don't in my lab don't don't make between sorting it will sort your array each column for itself and sort row where you can sort the row and the column you choose one and

notes

summary

39m 25s

1.3.25 Plot of the sorted data

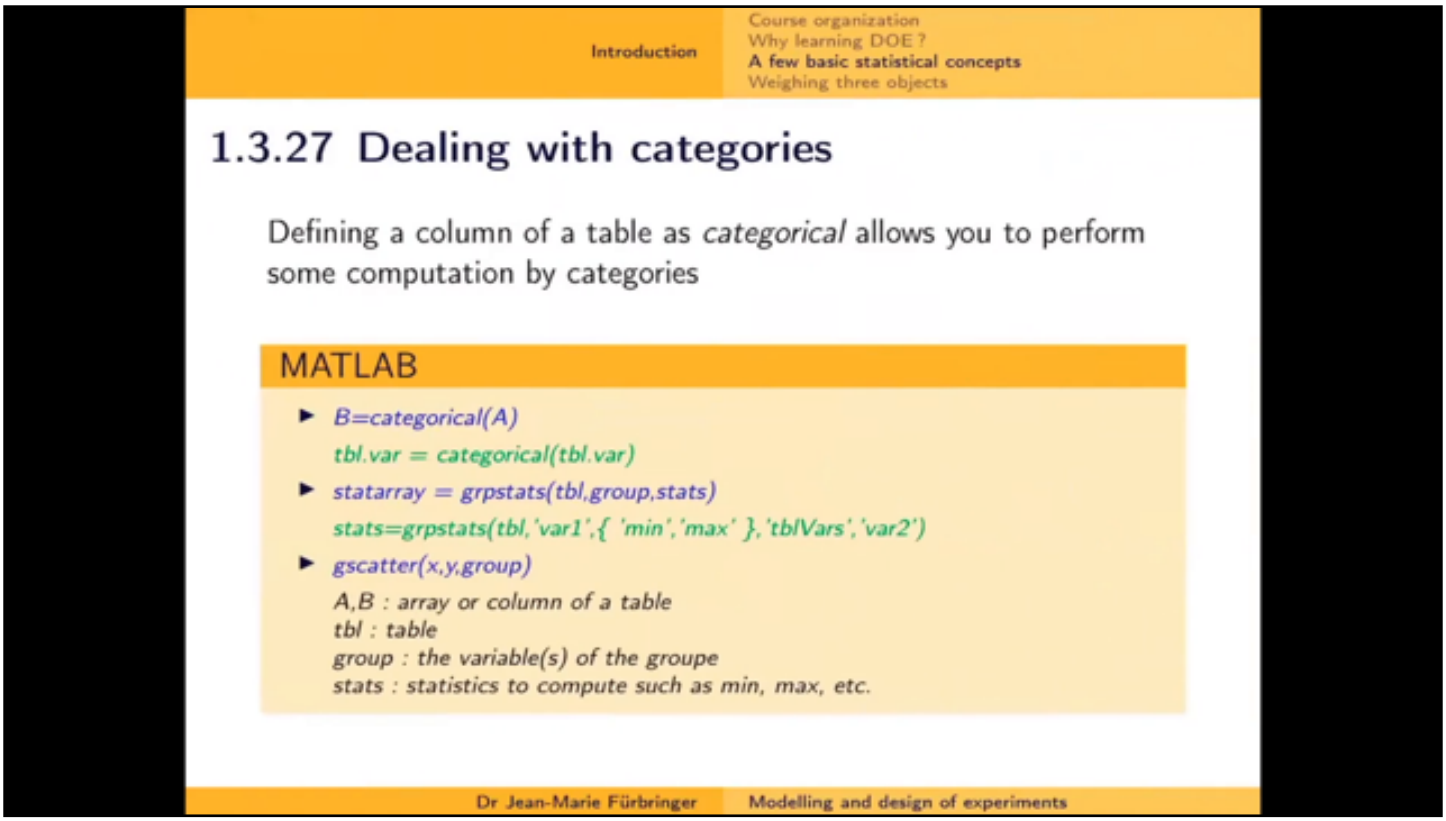


you sort all the other was with this one okay Pareto principle so it's also very important

notes

summary

40m 1s

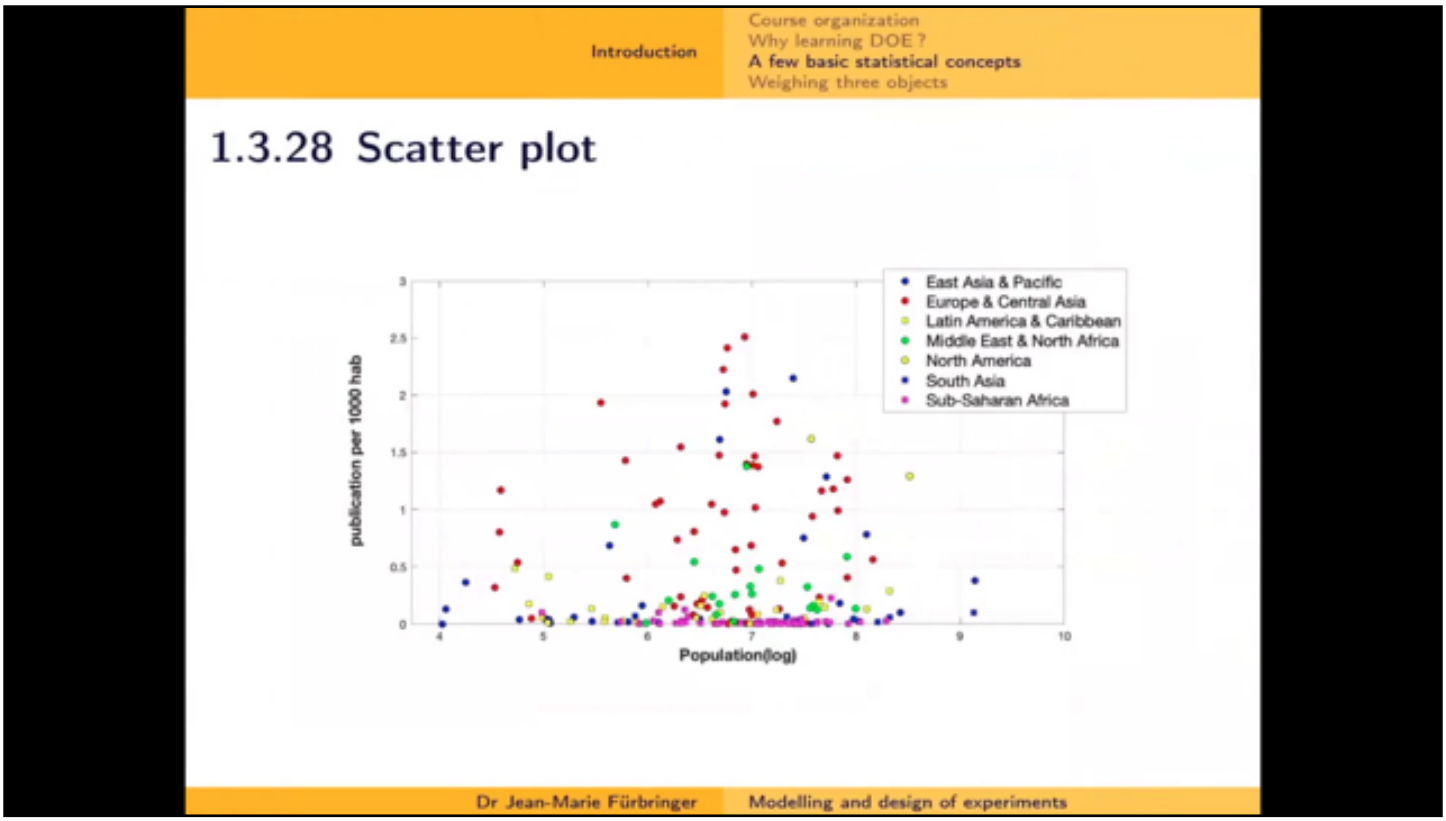


engineering to identify what are the minimum of factors that have the maximum of consequence so sometimes people say the 20 80 percent rule these ideas that have a 20 percent of the cause at 80 percent of the contiguous it's never exactly like that but sometimes it is presented like that what are the few causes that have most of the attack and so the idea is ordering your consequence from the one having the most consequences and one less consequence like that and after you can even add up the different participation and eventually you see where you get 80 percent of the consequence and how many effect and how many factors you have in 80 percent of the consequence and Pareto principle Pareto live in Los Angeles that's why we like it so it's why you have a liceo Pareto very important statistician too especially for the

notes

summary

40m 2s

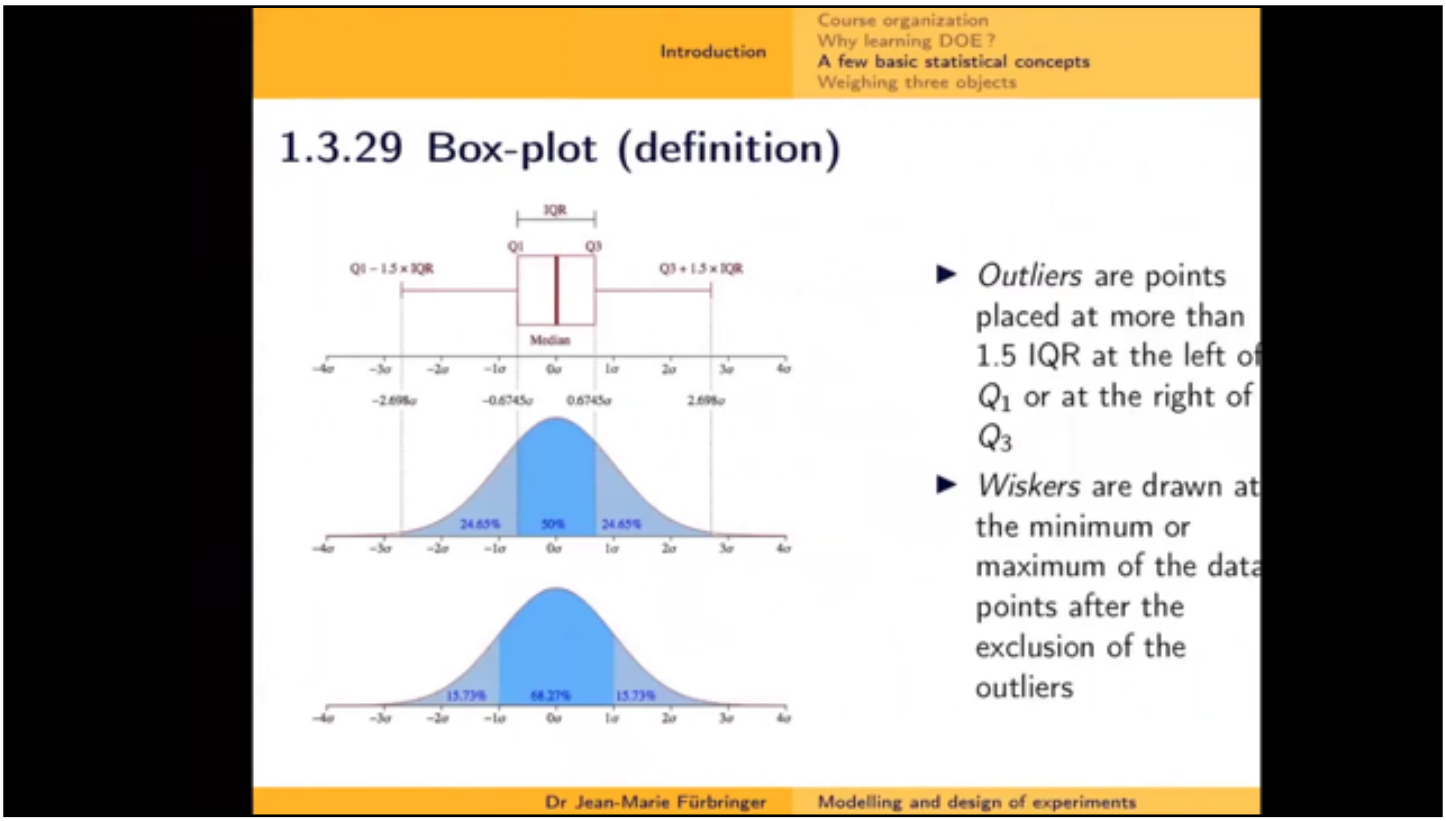


economy I would like to let you know that in my class and eventually in Python even if I don't know you have function for working in this category so you can define a variable as a category with the category command and after you can make group statistics group scatter it's helped you quite a lot for winning time when you are making this element you have to analyze that

notes

summary

41m 18s

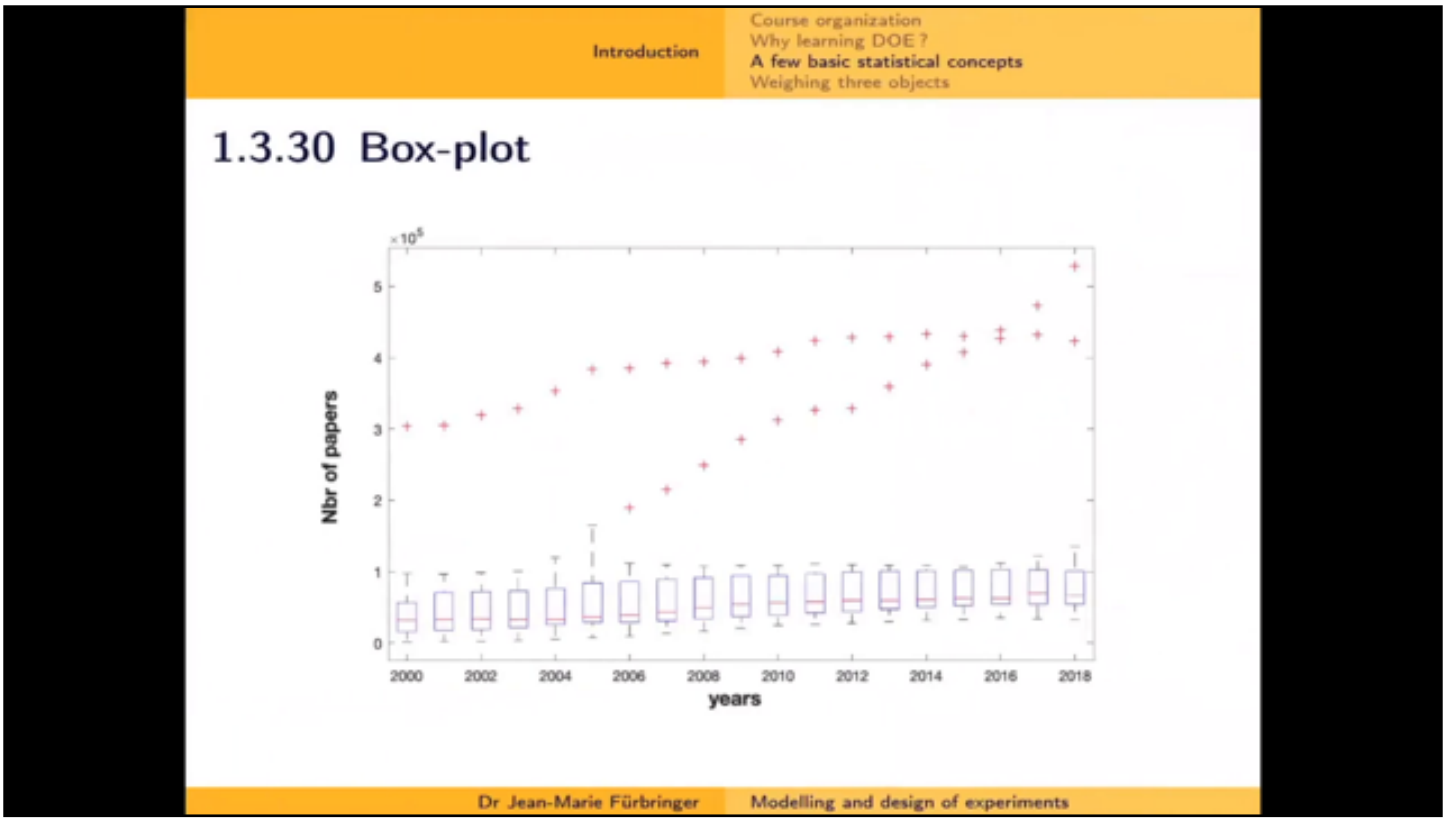


scatter plot are one of the basic elements for understanding that means that having two distributives that are was two two factors you can make it by groups this have been made by groups

notes

summary

41m 55s

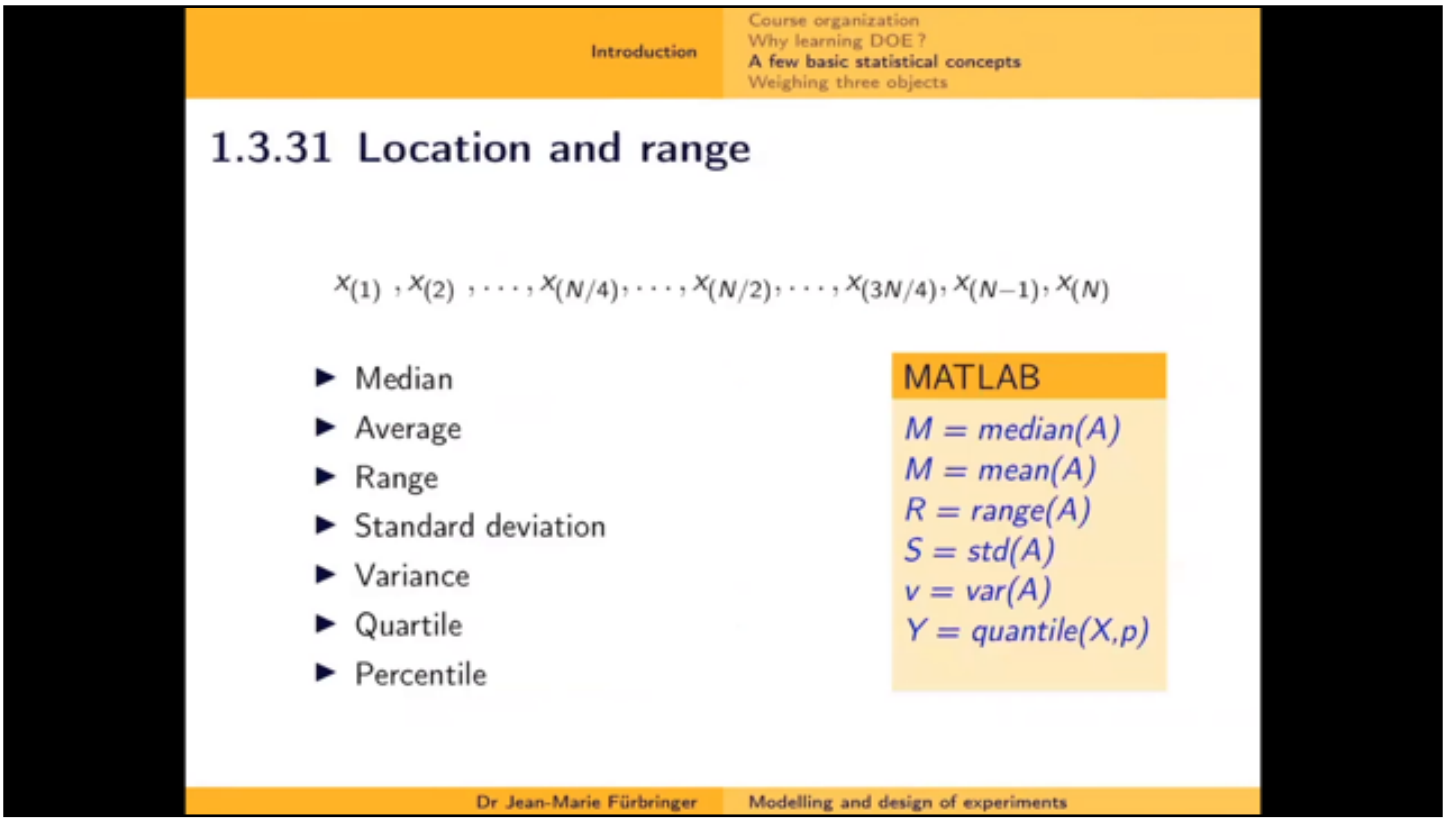


so very very practical box plot and provided it's not the first time you will learn about box plot is a way of summarizing several series and compare them you have in box plot you have the median so the value was the probability of 50 percent of the probability and right and left side after you divide your distribution in quarter of probability so around the median the left 20 percent probability left is 25 percent at right so it give you the box and after you have the moustache I don't know okay and so it's not exactly so it's the quartile one minus 1.5 the interquartile range so the interquartile range is the length of your box so it's not representing exactly the normal distribution because the normal distribution has the plus minus sigma is bigger than the quartile but it's close to so it's a very easy way to summarize when you have a lot of series and you want rapidly compare them box plot is a very interesting so it's also a way of this of the germinating outlays so if you are really data that are outside of this moustache we usually say that they are outlays

notes

summary

42m 14s



so this was for the publication a way of rapidly seeing what's happened and we see exceptions so exception was China and United States each time and you see the rest of the country which quite it's really interesting to see the evolution of things over time over a factor

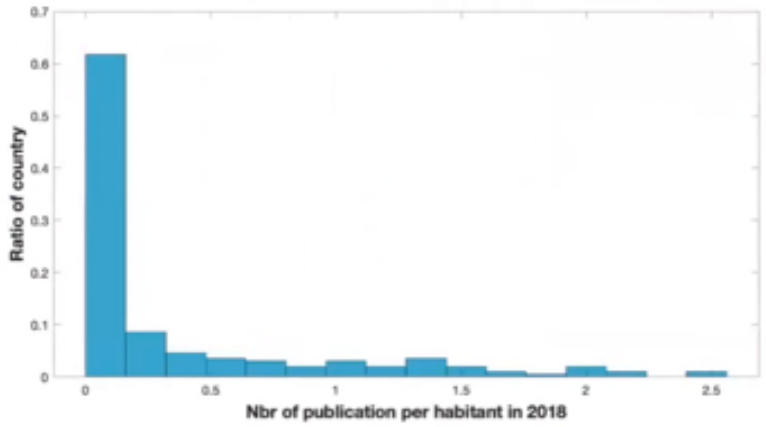
notes

summary

44m 6s

1.3.32 Quartiles and histogram

	Mean	Deviation	Variance	Q25	Median	Q75
2000	0.18	0.36	0.13	0.0025	0.018	0.11
2018	0.37	0.57	0.33	0.0128	0.08	0.48

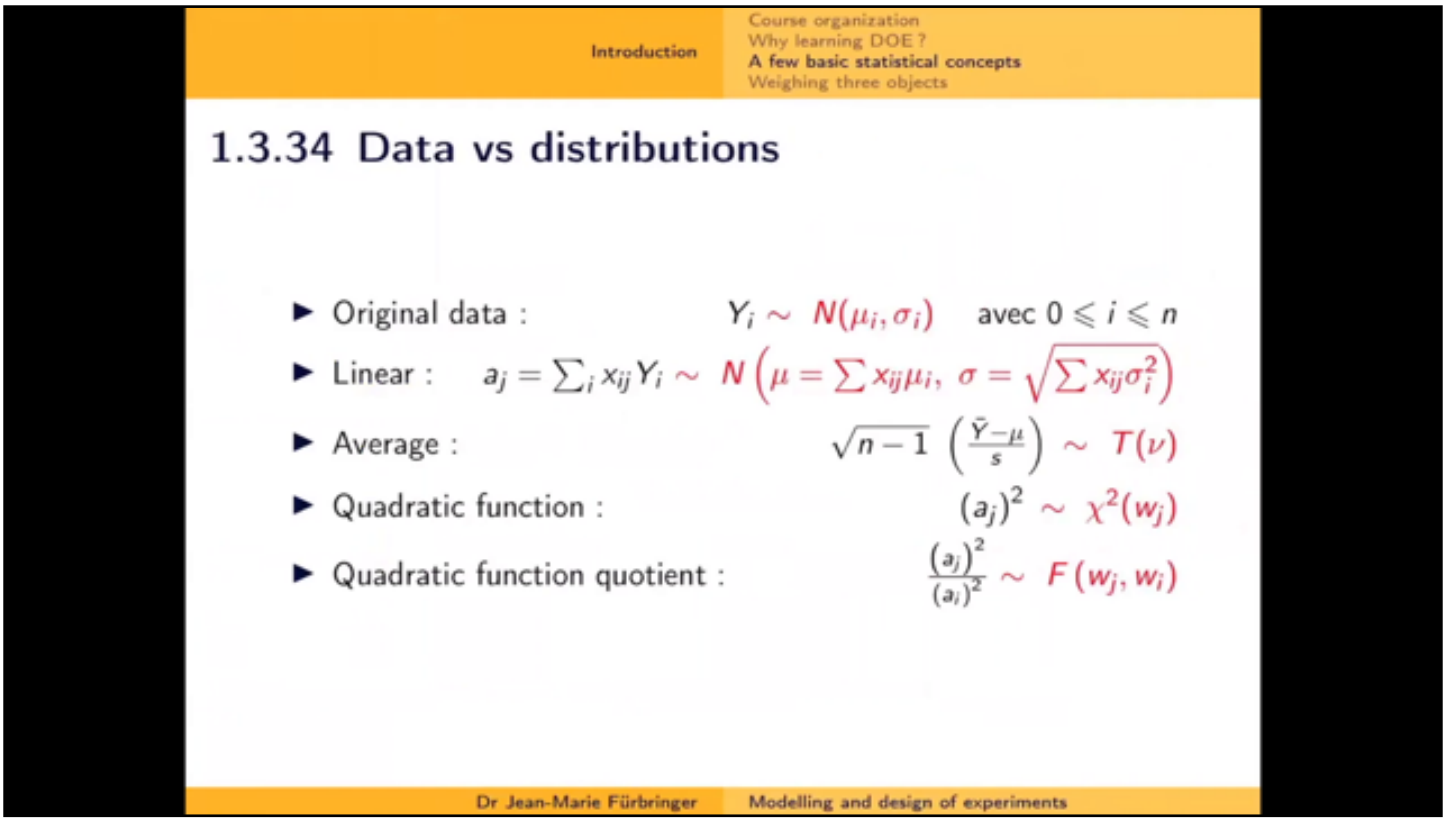


etc location and range so you have different function media mean range standard deviation quartile so just remembering you that MATLAB is doing that location and range a few calculation

notes

summary

44m 30s



I go rapidly and this quartile and histogram let you just describe what easily in first approximation distribution yeah this I will come back next I will you don't need that you don't need that for today so I will stop my course

notes

summary

44m 48s