Department. Statistics. Statistic suppose. Do you want to study? Don't say that. This study where at least four hours in. No, you are not very smart. You don't understand you for four hours. What will you do?

So you will go to your friends. You will ask your friend how many hours do you study in that day. 6% is only for two hours. One person says I was 31. Some other person. Do you have data? Sister we actually went to six students. Turning point. What will you do with this? This. Either you will have this data. 6. This. 4 by 6/10 plus 2/12/12, plus 2017, 17 + 118 and it is 28. And you have taken observations. Yeah, we're already seen it 64. 6. Actually you don't by 6. 6. 24 four remaining 6 miles and 36, six and 36, six, six are 36 packs on. 1.26 No, you cannot argue with your parents that you should. For more than. 4 hours they're reading because the data that you have collected actually comes out to be the English students studying for 4.66 hours.

So when you study statistics, you will call this. To suggest a demonstration, have not defining meaning. Suddenly you are still not happening,

So you will find another base. I will arrange this data and what is the least? Taste. The next phase is. That one is. 6. 6. I'll be using. What is the value of? Since our data of six elements, I will take four and five and I will take 4 plus 5/8. Equal to 1 nine plus. Still, you are getting that you should study for more than 4.5 hours,

So this kind of calculation. Ordered even if you order it in another way, like you start from 3:50. What is 4.5? Anything that you want to see is if there is any repetition that many people are study. Suppose if there is one data .65 introduction data points. There are two people who are studying for five hours. There are two people studying for five year olds,

So this is 5 is the number which is repeated most of the times. Particular thing. Space you cannot. Parents that you should study qualities than four hours during exam. We justify our claims when we study statistics that just gives you a free video of what we are doing in statistics. Only in these measures are not enough, but we will devise all majors, study statistics and data.

So let us. This was a motivating example. Let us know the actual term statistics.

So this statistics is used in two ways. One, it is a discipline in which we studied. Statistics which used to. It's just a list of data. Which you may you swing by the bottom depends on you,

So just stop it. Stop it.

So suppose I say. They want the book on educational statistics. Statistics in India? What does it mean? It means there is some which exists which will list all the data values which may not infer something about the data.

So it just gives you the list.

So in the UP. Did audit themselves. Australia. Constitution. Institute. Which is not. Institutions, it just gives you this data. About the data. If I make a statement that. 20 states. Statement.

So I study statistics of strong data to get better results for my investment. This statement means I actually collect the data and analyze the data, and when I analyze the data I get something which I use. I get some statistics which I use. Change my investments or to optimize my investments and

So this is the statistics is used as anticipated in which you study data set just recording of that data. One on educational statistics in area where you will not predict any. These days in our Times of India or any any news you get. Dear.

So this pacifies. Today it was moderate. Air pollution was moderate. It was worse. But this is only classified that way. It is forced. It is moderate and. Of the cities in which they. That is, you get consistent data about this. You can try to infer something about the city's health.

So all these things are nothing but examples of statistics,

So statistics actually can be divided into 2 parts. One is a discipline in which the study data and another one is stopped at the values. No. Definition of statistics. One word is common. That is, we have. Also, we have used the word. What does this data mean? There are many things that are coming in these days. These are data analysis data.

So what does this data mean? If you look at the scientific world, if you look at the dictionary meaning of data, it's nothing but prayer and of the world. And this data means. Piece of information. Piece of information.

So what is data? It's a plural of this. That means it's many pieces of information,

So there are several pieces of information which you will try to put it together when you

try to put it together in a meaningful way, you will get statistics. Statistics is nothing but the process. From which you then make some sense. Now that we have understood the world meaning of the world statistics, it is a process information which comes out of data or statistics is anticipated in which you studied it or otherwise its list of data values. Now let us try to study statistics

So anything that you study in statistics has to deal with.

So let us start with our first topic, that is collection.

So how may you connected? There are no ways in which collected one is primary. 2nd. Let me see what this primary. Let's see what this primary collection of data primary is. When you want to collect the data, you will actually go to each and every subject and collect it. For example, when I gave this example. Actually went. Friends and collected the values or seems to find one.

So this is the source of primary data. Right? Actually do this. What is secondary data. Instead you go on and. You see how many of us get steady? And you collect the data and try to refute the claim of your parents. That is interesting.

So it is something like you have not actually done the survey. But you have collected the data from the people who have already done this. 30 seconds. 2nd.

So it may contain some statistics. Persistent. Has tried to say. How the primary data is obtained from the secondary data? United State one. Suppose we want to contact the server. Fights.

So what we're gonna do? You can do it. You can do this your school database. We have a data set which has. He recorded the fights of the student. Or you can go to each and every candidate and try to figure out what is the height of that particular person. The other two ways in which they are. If you want to. That you are. Setting. Maybe 120 centimeters? One 5177. We will keep it short. So. Superstars and now I want to represent this data. Present this

So we have collected the data and you have actually seen the data set. What are you doing? Now I want to represent this statement formula represent. I want to see some salient features of nested

So first. As it is, we cannot create any sense out.

So what they do is they will just clear arrange the data set. What is the least element is it's 120. It is 125. Is 120.

So this is not so. I hope process this data now. No, if I if I ask a question, then what Rams the heights of the student? Question what is the range of the data? What would be the rate it should be?

So what is the least value? It's 120. 70s and 80s. Minus one should be the reefs.

So the students might release. This is the new information that is given to me out of the data,

So it's this is called statistics statistical. Statistics is so. Yes. Statistics. Now I want to represent this.

So for this is a small data set for how you need a larger data set. Let us say I will consider the mass of. The number of students. Class. These students.

So now you just see this problem. Considering the box of 20 students, 200 marks. These students out of class,

So these are the marks of the. While this data is difficult to summarize this data, we can if we want to reorder. Instead of this, what we can do is we can actually try to emulate the. The data. And then just say. If I order this data, how many parts are printed?

So let us try to. Oops. And the number of things in the box. This is part. The event happened. First, I see the marks. What else? One person is there 20 hours. The answers. Save the box. 636 How many people are in 36? What is the next number? How many people? Or people. How many days? Aspect. Second time. Yes. To see what is happening. This. 36 once. Twice. And what is that? Next 4 tips. 17 several details. 71

So only 72 is. 1st. Only once. Next

So let us see how many types, least once it does. Justice. Once. The reality. 88 years. One year twice. That what appears only there are two numbers that No 92.

So 92 are appearing twice. Once. Shut itself up to 13. That's what you have to do. 9 + 3 + 4. 24 426 + 420 minutes. Now you take the difference between these two representations. This representation. Presentation of this representation. Presentation of our data, which is started in a more disciplined manner. This is the point. Is really important because it gives you. Quick snapshot. Just. All this data. Quick snapshot of the data. There's

such a table is called. Distribution. I am because here I am using only one number. No. But this is a sample of 30 students. It can go further. It can be a sample of 100 students. It can be a sample of 200 students. In this case, the representation that is given here. This is the raw data is not very helpful. In that case we have to go for the distribution of this. This is a good representation. This also can be done. No, you have many people. Plants. 100 plants 1st. The procedure is done for hundreds 100 primes. The raw data. Because I. I cannot go and write all the 100 points.
So I have no choice.
So what I will do is. Just populate this two times. France. But we call it the 100 points. I need one. That if you want. So. What is the height of measure during if I try to record? You can go as many as one and I don't want to do that.
So what I wanna do is I will say the least, right? What is the least right in this?
So you start with. I will create a win, will create a class of data points. 2229 then I will create 30 to 30. 40 to 49. What is the maximum height? That's what I have to say. How far I can? You can see the number 80. Can I see the number which is better than mine? So 99. 60 to 60. 17 to 17. I want to figure out all those numbers. 20 to 20.
So it disgusts me. Similarly, you can do this. Similarly, you can complete this statement. 8 to 49. 60 to 6370 to 79. 20 it is disgusting. Now the total should be 137. Just. Our table. Now what we have seen here is. We are actually not dividing the table. One class where? Actually splitting it. Say for example, all the numbers that are between 20 to 29 are in one one class. Whereas when I did this. All the numbers are not in one class, they are individual numbers. And that is why we call this as frequency improvement. The name of this. Also. You are not interested. No. Here are some interesting observations suppose. I want to ask a question for how many schools more help if people sent out the plants survive, how many answer that question? Just look at this. Beyond that, all the plants that have survived. Basically we are that all the brands that are better than or equal. Just some total this. You will get. 71 has the time.
So you can easily see that 71. 71 schools percent. Parts. Interesting.
So now we are actually trying to do something and we are actually trying to fix something. 71 sports survive. You may further split this table if you want more listings and in the small you can further split this table from 20 to 25. 2016 29 and you can again. Morning. If you want the final of the data sets. This is the one. All these plants and all all the data that falls considered in integers. Suppose you have been given another. Have been given rights in. Do you want to come? You want to compute the data that is now the raw data. Somebody has done the work for us and they have given us. Table is .6 G. Wait here, we will continue and the number of students. These are something like this. 36 What do you want? What do you want? To make. 4650 these six. 61 to 65. Sorry. Students. But no, if some new data comes. So. And the person is sent 35 points. That is the right now I according to this team.
So it's a problem for that. 25.5 cases. Where will I want that?
So if I put it 35.5, I cannot put it in the pane of 36 to 40 I cannot put.
So there is a problem,
So I need to modify something about the net.
So what should I do
So the easiest variation that I can do is I can modify the class sizes accommodate these weights. Modify this process. Because I am considering all the numbers are integers and I'm having this problem if that means essentially having considering discredited. The data that I am getting is not displayed. It's a continuous data set. The numbers which are in if I have a display data then there is no problem with the division of 1. But now because I have a continuous data, I have to recorded this data by making some substitute arrangement. Just try to see. 3125 is that it doesn't, and the next successive column is 36 to 40. How can I accommodate the weight of 35.51 of these?
So what I can do is. You just see. This class. 6. Of this class and the limit of this class. See if I can do something like 25 by 26. Division of 25 point. No. This. What do you like can do for every successive class? I get the. 3130.5 five point. But this should not alter the classmates of other classes. What is the class 1st? Let's take the difference between the upper bound and the lower. Certainly to find the class with. Classmates. On class which is. What about you of the class? Wireless mouse.
So for example, if I consider the class. 31 to 35. This is the class, but it will be 35 - 31, which is equal to. This is the password for. Similarly, we find another class which is 36 to 14. 36 - 1 is 4040. Just 36, which is a good way for being one. It will be the

class. If I change my classmates. Then I should do this uniformly for all other classes. It's 35.5. Minus 30.1 that means I am changing my password to make one. This class might equally fine. Change should be reflected in all other classes. Just change the class. 35 we changed. 25 36 to 40. Should be changed. 5.5. What do you want? Our next class was 41 to 14. Thank you very much. Point minus. 351 All these will have class. 5.5. All these datasets,

So let me. With these two additions. I have added the two points one and they are value 100 value of 35 point $5.5 million one. Yeah. This new modification. 45 point. 51 Now we have a coordinated. Some time.

So the 25.5. Actually be included in this class. Place this number. 6. 5.5. 45.5 this particular. 35.5 is the first forty point. Five points. Is 40

So 45.5? This is the force. 3 + 1. Now the tally doesn't change.

So this is all we accommodated in the country which was not earlier possible because of the discretization. But now we make the data to be continuous. Now you give me any number in this range they can accommodate that. It's the small things that we have that we have where the collection continues. So.

So we have represented the data in the tabular form. Now what about the? I wouldn't description of the data in the graphical form,

So that is now come to another. Because representation. And we have. Many options for the first one. Autograph or bar? Modern. Report you don't have any numbers associated with your data. One example. Let us take one example. Amount of rain. They are. January. Someone? 4. Basically everyone.

So now this is the data available. I want to compare the data set,

So I want to actually see any graphical. How do you look at the data? Wanting sex. No. Look at this. You don't really need any success or scale. What you need is just. Something like this here? Possible rainfall is. They are saying in the month of January. What is the? One in January. Just go. It was. That is the case. Water. There's that. Process. It's not a number. For example, the data set can be how many peoples languages? So somebody might say English Spanish. Chinese. What are the Hindi? You can actually draw on each of the languages. In such cases your part. These points. X axis understand why axis is nothing but. Whatever they are. From this bar chart, what you can see is even those who are visual readers. It is more visible than in the one part that we had more. In the chat. We've had force and in the month of July we had at least. With representation, which is not very clear from this data, the tabular data is more clear when you actually graphically represent the. Some more. Program. Thank.