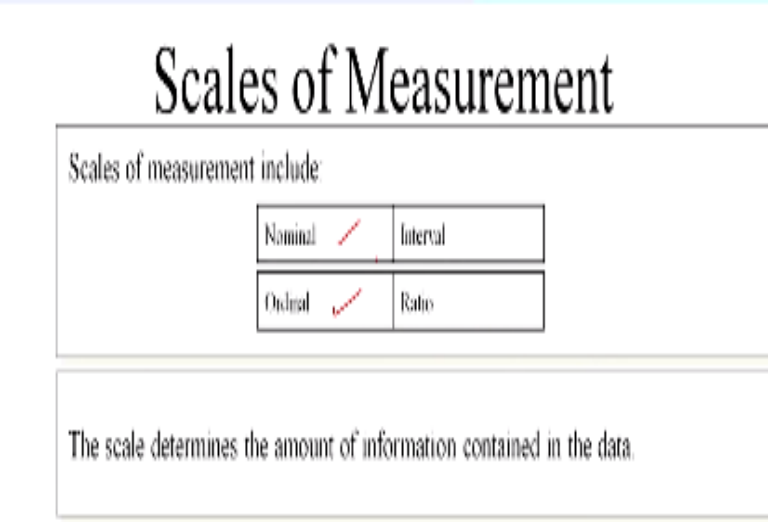
Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology- Roorkee

Lecture – 03
Scales of Measurement

Good after noon friends, I welcome you all in this session. As you are aware in previous session we were looking at 2 different types of scales of measurement and we have seen that basically there are 2 types of scales numeric and non-numeric. So you can say that non-numeric scales are nominal and Ordinal.

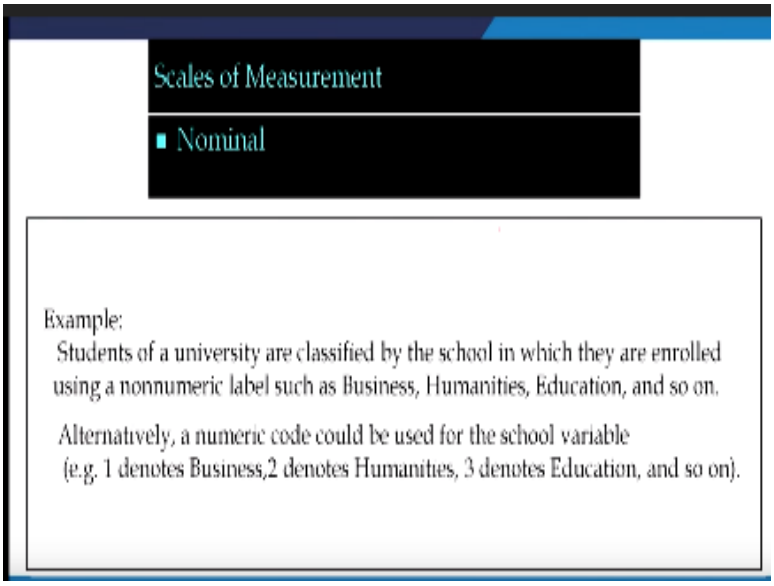
(Refer Slide Time: 00:48)



And numeric scales are interval and ratio. We have seen in detail in previous class about nominal scale and the purpose of nominal scale is to identify a person a product that a group or any other thing just for identification purpose. For example, roll number of a student it is just for the identification of a particular student. It does not mean that roll number 1 and roll number 30 students then roll number then the student who has got roll number 30 is 30 times intelligent.

So they are just for identification purposes. So we have seen an example as well in previous class. So let us say you can have a nominal scale like.

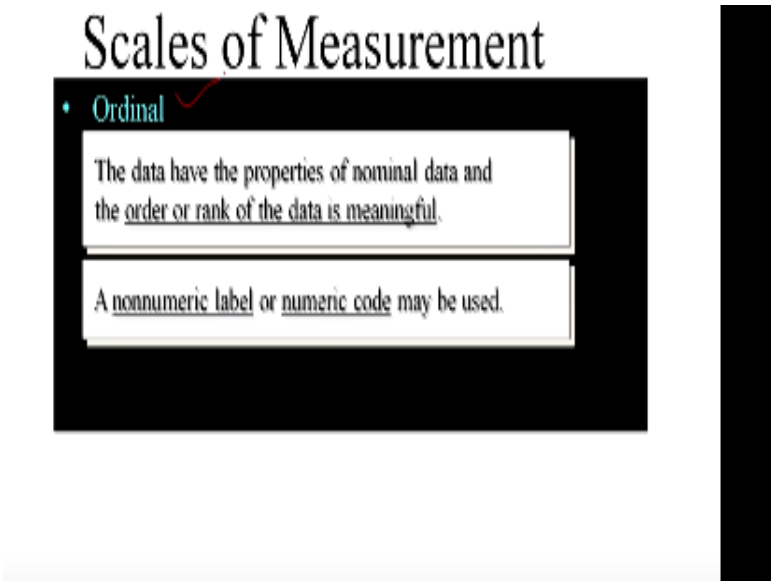
(Refer Slide Time: 01:36)



You can have a different departments in the university business school humanities, department education, department civil, electrical, mechanical. So these are nothing but nominal scales for example gender is also a kind of nominal scale. So you can say male female group 1 group 2 yes no right? All these are nominal scales and of course you can convert nominal scale into a metric scale by assigning some numbers to them right?

For example, 1 to business school 2 to humanities and so on right? Let us look at ordinal scale. As I said nominal scale is just for identification. Ordinal scale is basically a scale which tells you that this particular person has got more characteristics than some other person or some other persons are a product A has got more characteristics than product B. So this order can be measured through ordinal scale.

(Refer Slide Time: 02:54)

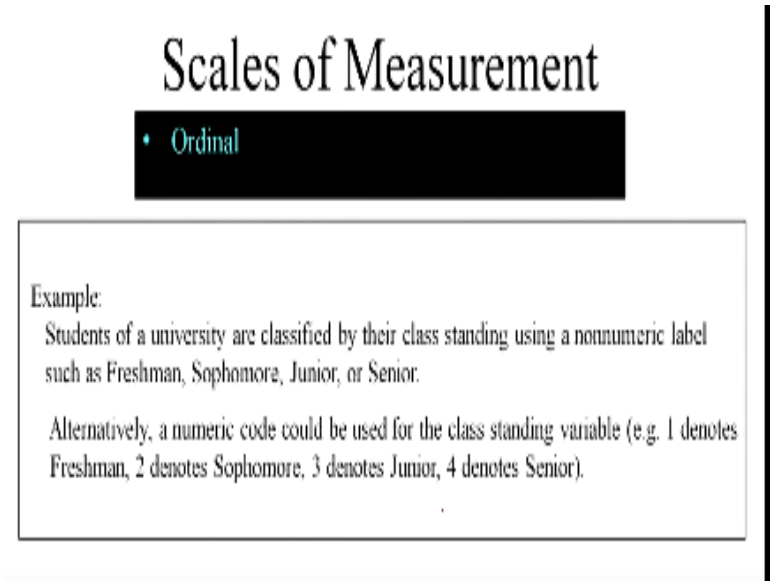


I will tell you a very good example let us say in an examination a 100 students have appeared right and all of then wrote 5 papers right? And each paper is of 100 marks so total marks are 500 is it not total marks are 500. Let us say the topper of the examination is let us say scored 495 marks. So we will say that this fellow got rank 1 so this is nothing what rank order scale 1. Then you will have some more students.

So similarly let us say this is serial number 1 right? So you will have total 100 students is it not? So the last student will have a rank 100 so let us say the last fellow got just 300 out of 500. Now the fellow who is at second rank it means this fellow has got more features or more characteristics then this fellow. But rank order does not tell you how much is the difference between a rank 1 and rank 2 students as far as marks are concerned.

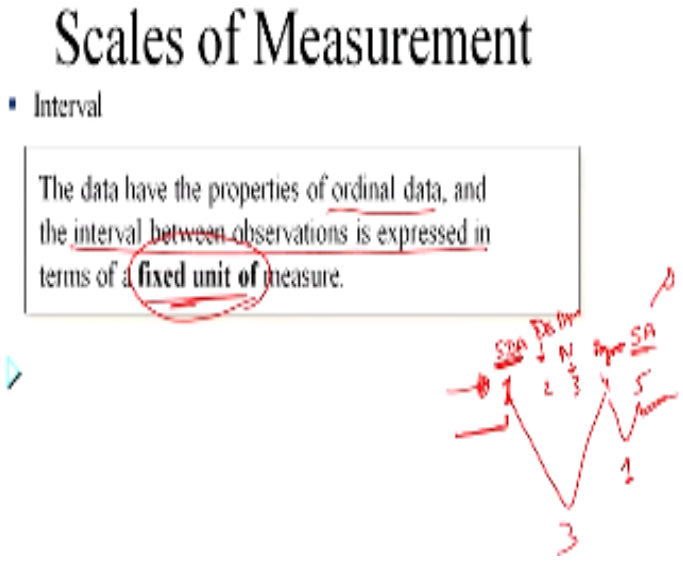
So that is the major drawback of ordinary scale. It will tell you that this fellow has got more marks is better student but how much better is it not? So that that is the drawback of ordinary scale. And this drawback is it can be removed if we move on to the next level of scale or the next type of scale which is interval scale. And interval scale is a metric scale. Let us look at example of ordinary scale.

(Refer Slide Time: 05:13)



So students of a university are classified their class standing using non-metric level so such as freshman, junior, senior and so on. So you can just you know give them numeric code 1 denotes freshman 2 denotes this 3 denotes junior and 4 this senior student. So this rank 1, 2, 3, 4 right. So this is rank order scales example. Let us look at interval scale so interval scale will tell you how much more characteristics. A student has got compared to other students which was not possible in case of a rank order scale.

(Refer Slide Time: 06:04)



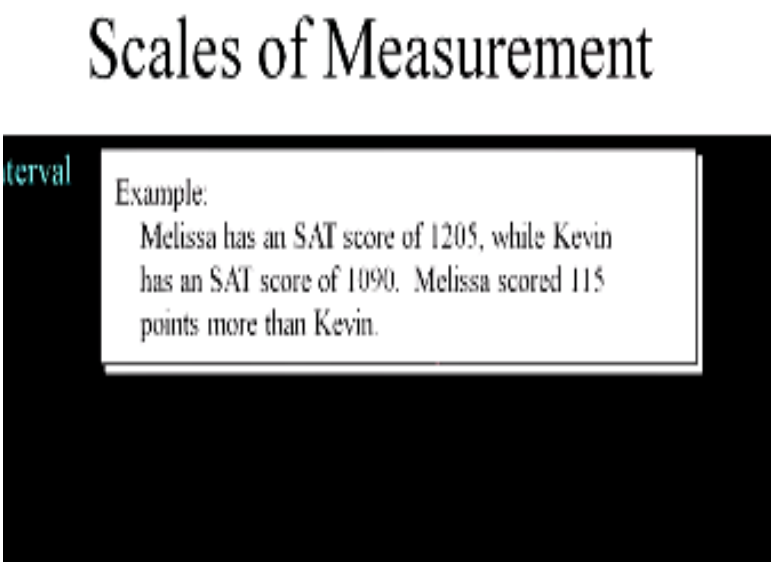
So in case of interval scale the data how the properties of ordinal data of course and interval between observation is expressed in terms of fixed unit of measure. So let us take an example of interval scale so let us say you are measuring attitude of a person towards a new brand of car. So

you can have a scale like this 1 2 3 4 5 a Likert scale having 5 categories. So if the respondent is strongly liking that particular brand of car so he will take over here.

So he is strong liking and very less liking right? Or you can put it in in terms of whether the fellow is strongly agree or strongly disagree right. So you can have scale like this. So strongly agree this is neutral right? this is strongly disagree. So this is somewhat let us say disagree these are 5 categories of the scale. So we know that the distance between these 2 is 1. And we know that distance between these 2 is 3.

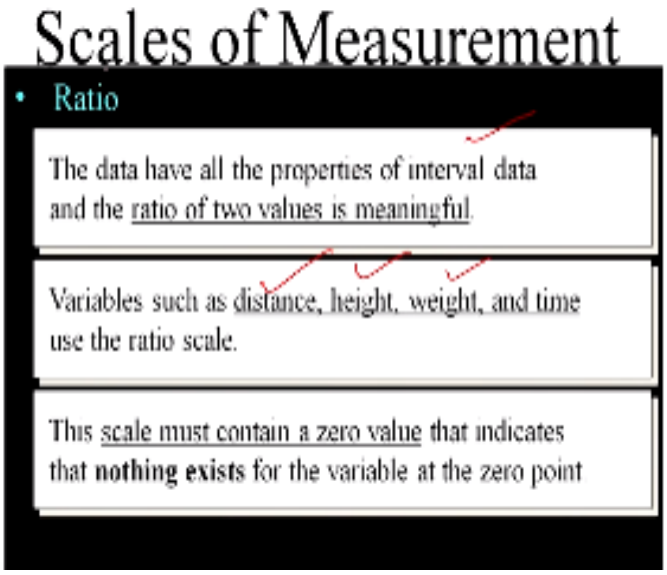
So this is we can have fixed unit of measure right? Because there is something fixed over here right? So if you take this as the starting point then this the last point1 you can take the vice versa also. You can have this strongly agree as 1 and strongly disagree as 5. You can have a scale like this as well. So this is interval scale it tells you how much characteristics a product or a person has got more or less compared to others. So they are always numeric let us look at an example.

(Refer Slide Time: 08:20)



Melissa has an SAT score of 1205 while Kevin has a score of 1090. So Melisa has got these many more marks or Melisa score is 115 points more than Kelvins. So this is so you can know how much is the difference right? So here difference is 115 this is interval scale. And finally the ratio scale which is which is again a metric type of scale.

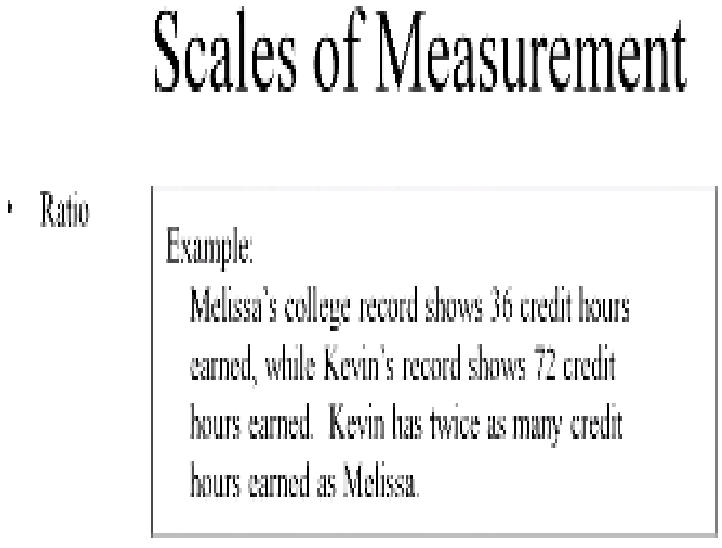
(Refer Slide Time: 08:58)



The data have all the properties of interval scale and the ratio of 2 values is meaningful in case of ratio scale. So we measure you know variables like distance. So distance, height, weight, income, number of houses let us say number of family members. So all these are nothing but ratio scales because in ratio scale your 0 point is fixed. So when we say 0 while using a ratio scale it means nothing existed for us.

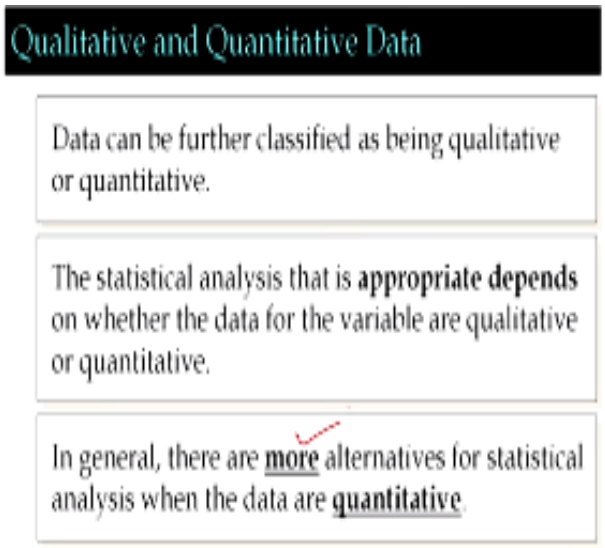
That particular variable is concerned when say height of a person is 0 it means of course you cannot have a situation like this right? So if these scale contains a 0 value it means nothing exists in that particular product right.

(Refer Slide Time: 09:58)



So Melissa college record shows 30 credit hours earned while Kelvins record shows 72 credit hours earned right? So we can clearly say that this fellow's credits are more than are exactly double then this fellow's credit hours right? So this ratio is important in this case.

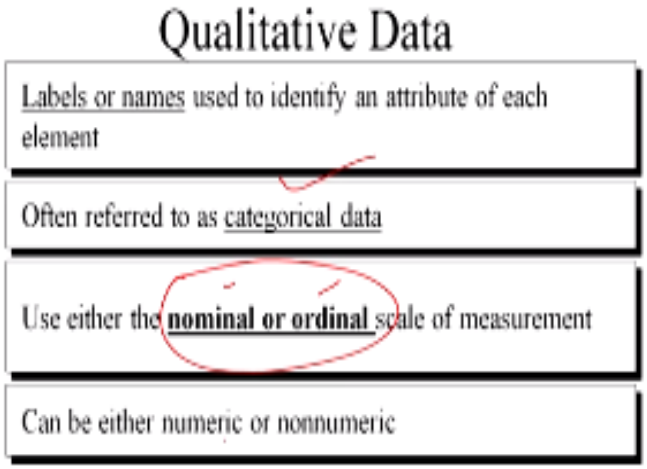
(Refer Slide Time: 10:28)



So you can have data again in terms of qualitative data and quantitative data. Earlier sessions we have said that you can in primary source of data secondary source of data. So you can have another classification of data qualitative and quantitative. Now what kind of analysis you are going to do on data it is very important to use a particular type of scale. So let us say if you are have got qualitative scale.

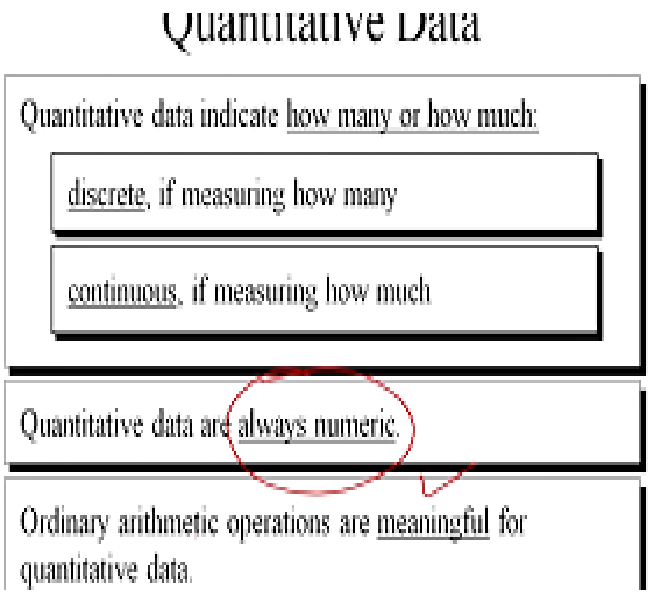
For example, a nominal and ordinal scale then you are doing some sort of statistical analysis. You cannot do some other statistical analysis on qualitative data. Similarly, if you have got quantitative data. Then also you have got some statistical analysis. So if you have got let us say quantitative data then you can have more options as far as statistical analysis are concerned and these options get restricted when the data are qualitative in nature.

(Refer Slide Time: 11:44)



So qualitative data are obtained from which two scales nominal and ordinal right? So this is also referred as categorical scale or non-metric scale. So use either nominal ordinal scale as I have already said. So again of course you can convert nominal and ordinal scale into some numeric numbers write. Some numbers can be given to these 2 scales. As I said some the statistical analysis are limited in case of qualitative data.

(Refer Slide Time: 12:27)

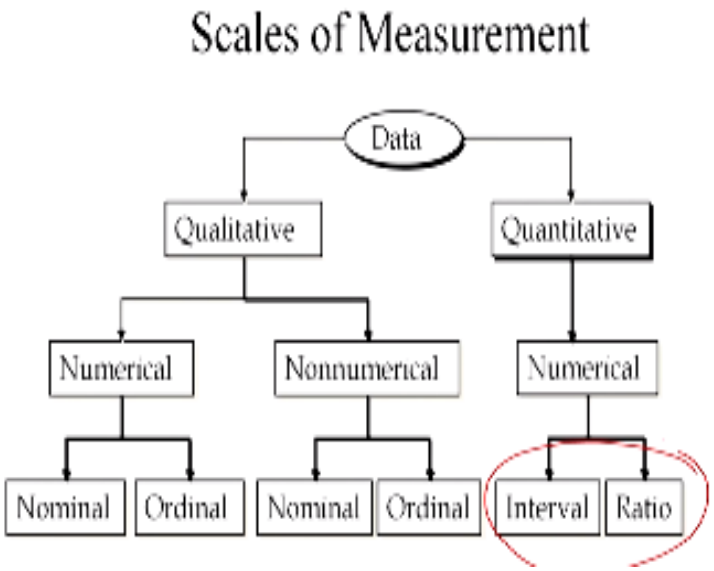


If you look at quantitative data of course they always have to be numeric is it not? Data again quantitative can be discrete or continuous. We have already seen discrete data and continuous data. Ordinary arithmetic operations are meaningful for quantitative data. You cannot find out let

us say mean of qualitative data right? Is it not? For example, red and blue what is the mean? You cannot have mean is it not?

So in case of quantitative data you will have some meaningful arithmetic or statistical operations right? So broadly what we have said you have got qualitative data and quantitative data.

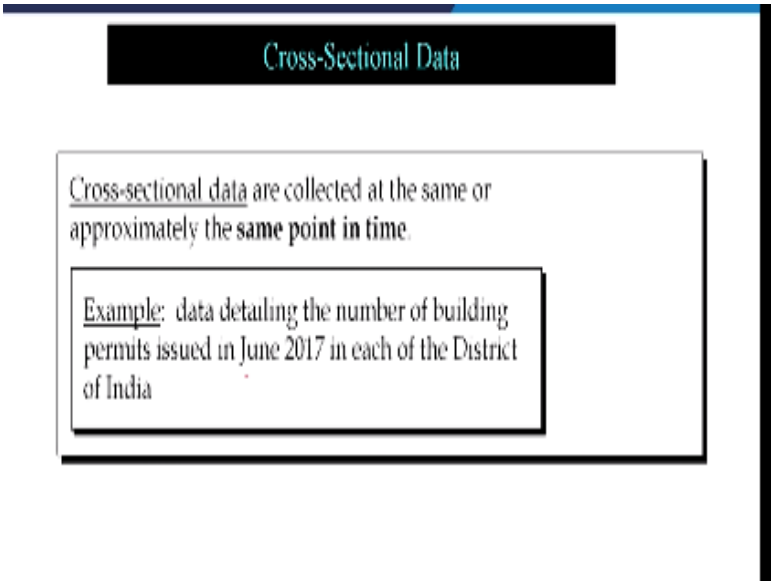
(Refer Slide Time: 13:26)



So quantitative scales interval and ratio scales. If you have got qualitative data you have just to write nominal and ordinal. Now this can be again numeral and non-numeral right this is what we have already seen. Let us look at what is cross sectional data? Cross sectional data means data collected over a period of time. Just one time you collected data. For example, the census data government of India collects after every 10 years right?

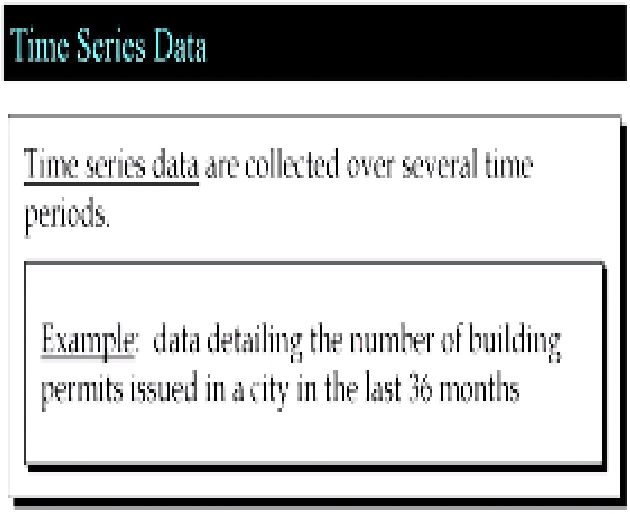
So let us say last time the data were collected in 2011 okay? So we will say that is just a cross sectional data. It is just like a picture of a situation right?

(Refer Slide Time: 14:22)



So example can be the number of building permits issued in June 2017 in each district of India right? And will have a cross sectional data right? But if you collect these data after let us say every 10 year if you have got data from all districts after every 10 year then it becomes a time series data right?

(Refer Slide Time: 14:52)

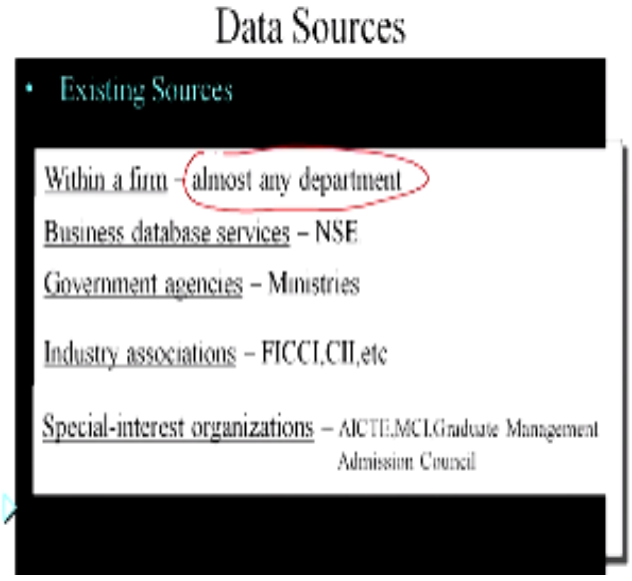


So data collected over a period of time are series data right or time series data rather than cross sectional data. So let us say data detailing the number of building permits issued in city in the last 36 months. So you will have 36 data points right is it not? 36 data points so this is time series data. So we will say we can say that time series data like video right? Or film of some situation right while cross-sectional data is just a picture right? There are different sources.

There are several sources within a firm. In a firm you will always have data in different departments or different verticals of the firm and be it related to let us say marketing or sales data or let us say a number of new products you have rolled out in HR department the number of people working in the organization how many people are going on leave what is you know the turnover right of the employees.

In finance data in finance department you can have data related to what are the sources of funds what is your return on investment, what is your return on asset and so on right? So within a firm you can have data right?

(Refer Slide Time: 16:19)



And apart from within firms you can have other sources like business database services. So you have got let us say in India you have got national stock exchange right? Or Bombay Stock Exchange let us say New York Stock Exchange right? If you talk about business and data based services so you have got let us say some other stock exchange of some other country right? You can have government sources of course government departments.

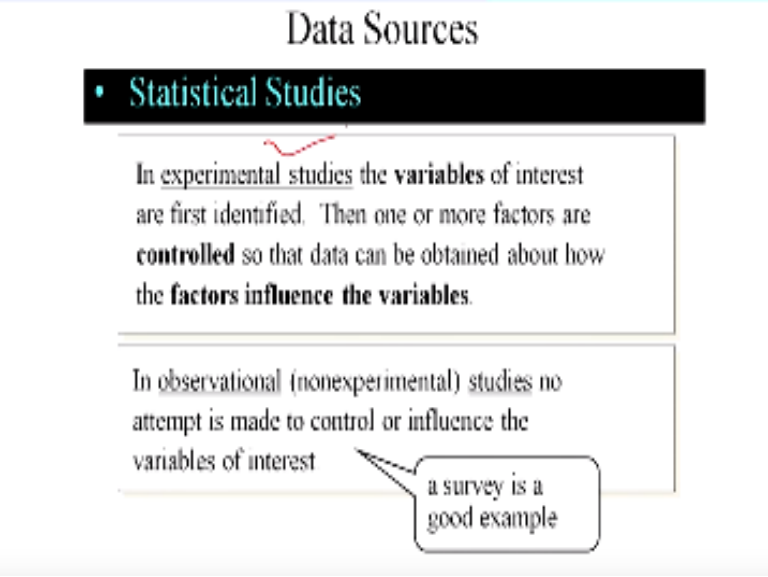
In India you can have different data from different ministries right? Finance ministry, HRD ministry, Ministry of Statistics and planning implementation ministry of water resources and so on. And so there are several ministries so you can get data from government agencies as well.

There are several industry associations. For example, in India we have got FICCI, it is Federation of Indian Chambers of Commerce.

You have got confederation of Indian Industry CII you can have a let us say SIAM society of Indian automobile manufacturers you can have data from NASSCOM right. So SIAM is there FICCI, CII all these are industry associations. National Association of software companies right NASSCOM. So special interest organisation so there are several other sources apart from these sources.

So you have got let us say how many institutions are working in India as far as technical institutions are concerned AICTE All India council for technical education medical council of India, pharmacy council of India and so on right UGC and so on. So these are special interest organizations from where you can get data. And of course Internet is one of the sources of data.

(Refer Slide Time: 18:25)

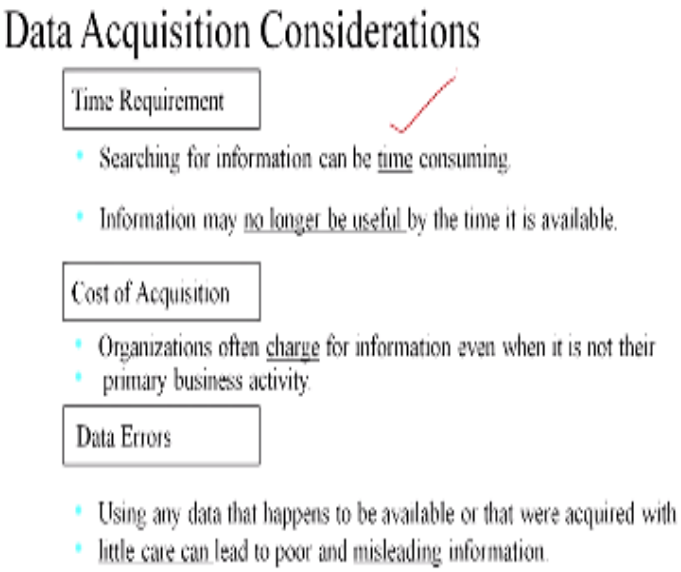


You can also data from experimental studies. So let us say you are doing some experiment in a laboratory. So you want to measure some output with some given input. So this is your output and these are different inputs in an experiment. So you can measure output right? So what are different inputs how those inputs are affecting output right? So these are experimental data so you can control some of these inputs right.

And you can see how this output is getting affected when you control or when you change one of the inputs or all of the inputs right. So this is known as experimental studies. You can have observational studies so here you are just observing some situation and collecting data. For example, as a survey through let us say survey can be put under this category. Okay so these two can be termed as statistical studies right.

Data acquisition considerations you need to consider so many things because you need to spend time you need to spend lots of money; you need to you would be taking time of respondents whenever you are doing any survey. So there are different costs involved in it so as I said its consuming process.

(Refer Slide Time: 20:15)

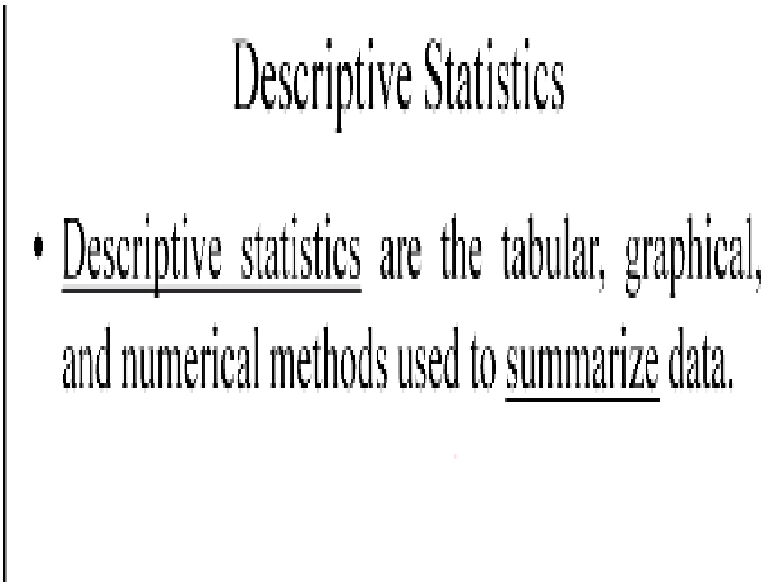


Time consuming and many times sometimes and what happens when you need some data and if data is not available then you will not be able to complete your studies. But sometimes what happens by the time you get data the data may no longer be of useful for your study right? So time is important factor cost of acquisition. Of course there are lots of costs involved in data collection.

So you have got field workers. Your supervisor sends to whom you are paying their salaries and wages. So cost of acquisition sometimes you acquire data from some other companies. So there are some syndicated syndicate services. So these are the companies which collect data and create

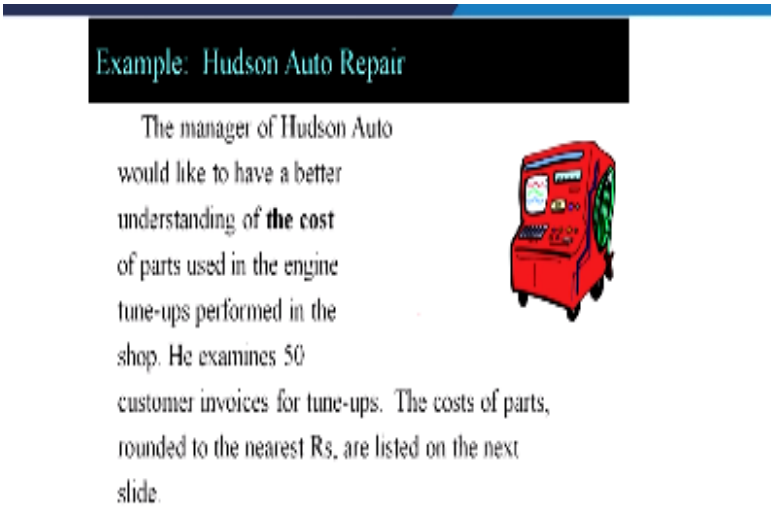
pool of data. So if you want data from that company you can get data by paying some charges right? So cost of acquisition is there and of course data errors try to minimize data errors.

(Refer Slide Time: 21:28)



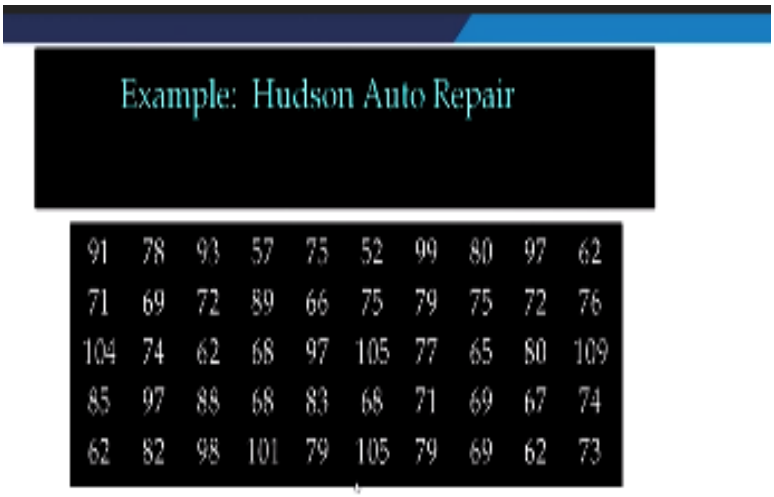
Okay let us look at data representation tools first we will look at some tools, tables, graphs for descriptive statistics.

(Refer Slide Time: 21:43)



So let us look at this example so there is a manager of Hudson Auto and would you like to have a better understanding of the cost of parts used in the engine tune-ups performed in the shop? He examines customer's invoices of 50 customer's invoices for tune-ups. And he wants to know what is the average cost of tune-ups. So these are the different costs.

(Refer Slide Time: 22:26)



So for let us say for 1 automobile this was the total cost for 50th this was the cost. So he had lots of invoices and he chose only 50 out of those lots of invoices. So let us look at tabular summary.

(Refer Slide Time: 22:50)

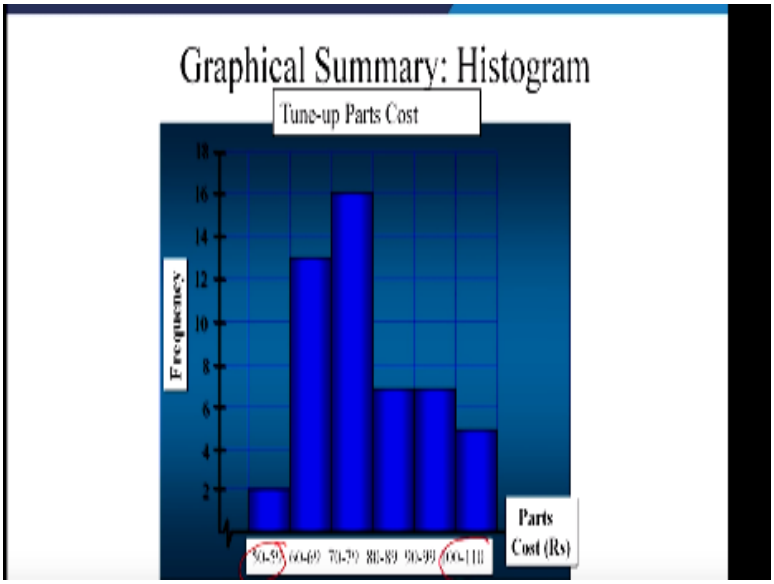
Tabular Summary:
Frequency and Percent Frequency

Parts Cost (Rs)	Parts Frequency	Percent Frequency
50-59	2	4
60-69	13	26
70-79	16	32
80-89	7	14
90-99	7	14
100-109	5	10
	50	100

Frequency and percentage frequency. So this is how you can prepare frequency and percentage frequency. So first of all what you need to do you need to prepare different classes. How many parts have occurred this many this much cost? So well say that there were only 2 values in this range 50 to 59 let us check. Whether there were 2 or not 50 and 59 right so 50 and 59 so this is the one and 50 and 59 yeah this was the second one right.

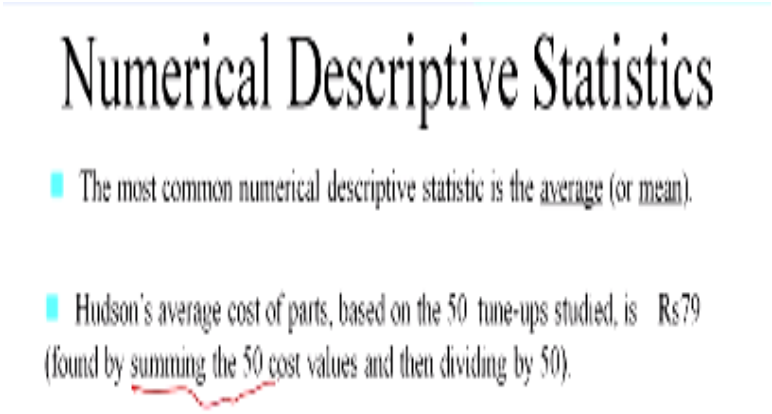
So frequency between 50 and 59 is 2 right. Similarly, you can write down frequencies over here in these classes right? So this total frequency should be 50 because you have taken a sample of 50 units calculate percentage frequency just divide 2/50 is it not multiplied by 2. So this is how you will get percentage frequency and this total should be 100 right.

(Refer Slide Time: 24:21)



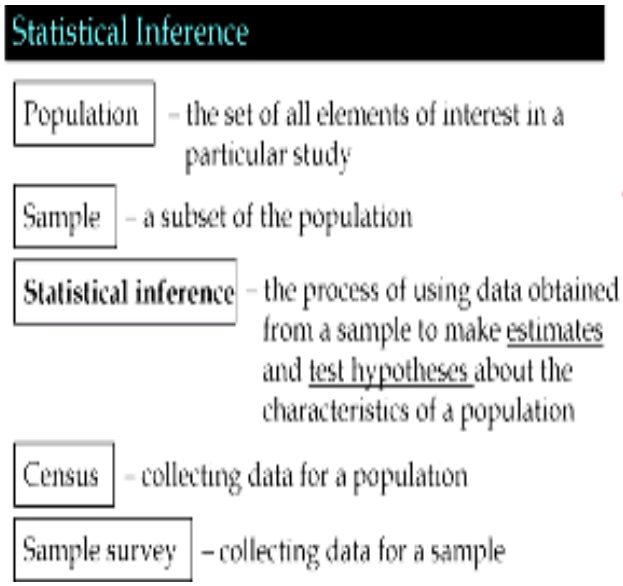
Now you can have a histogram of this example so this is your histogram these are your classes and these are frequencies right? So just by looking at this particular histogram one can get idea what would be the tune-up cost? So what do you think in in which arrange this cost varies is it not here right maximum is 16 right? So this is so the cost should be somewhere between 70 to 79 so what would be the exact average cost? Of course, you can just take the mean of all those 50 points.

(Refer Slide Time: 25:04)



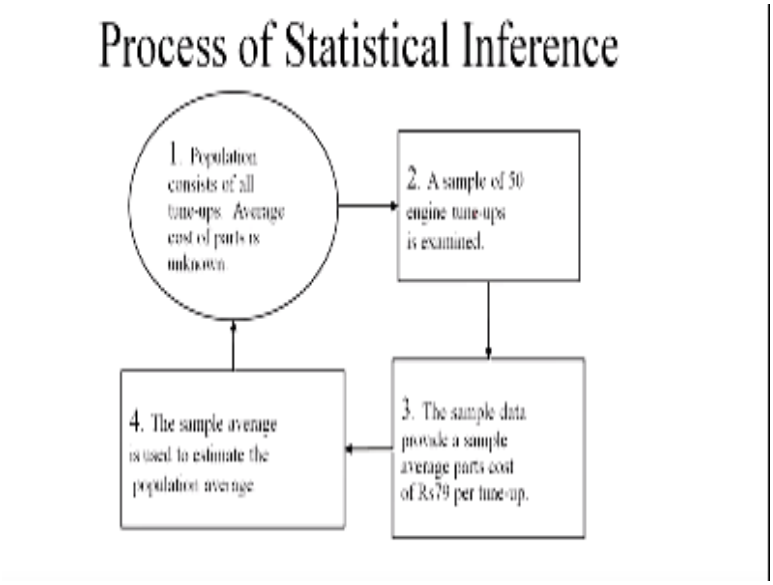
So you just take you just some summation of all these 15 costs and divide them by 50 right? So it is 79 rupees right? So this is how you can prepare double frequency and frequency percentage.

(Refer Slide Time: 25:34)



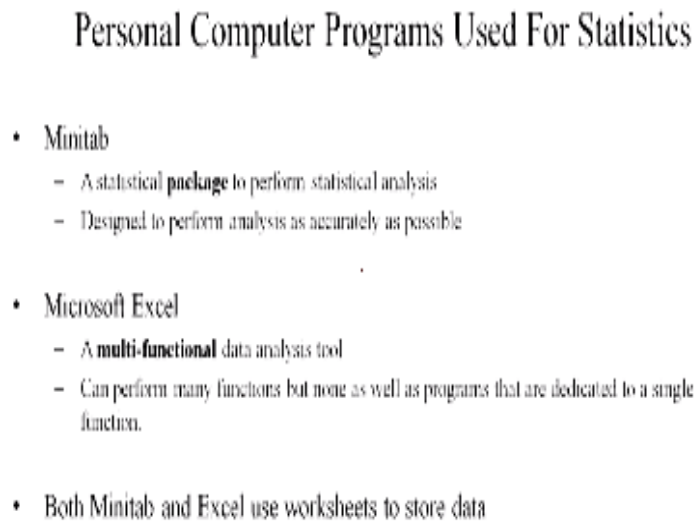
So in this case what we did in in previous example we did infer something from statistical analysis. So we will call it a statistical inference so there was a population right all the cars which came to that particular auto garage auto shop we took a sample of 50 cars and on the basis of those 50 cars we estimated and we concluded that the average cost was 59.

(Refer Slide Time: 26:19)



Process of statistical inference what we did so there were several units in the population. We took 50 we just took the average of all those 50. And on the basis of data average of the sample we have said that the cost of repairing each automobile was 50 rupees for all the automobiles which were there in population right. So what we have done we have infer something about population from sample is it not? So this is the process of the statistical inference.

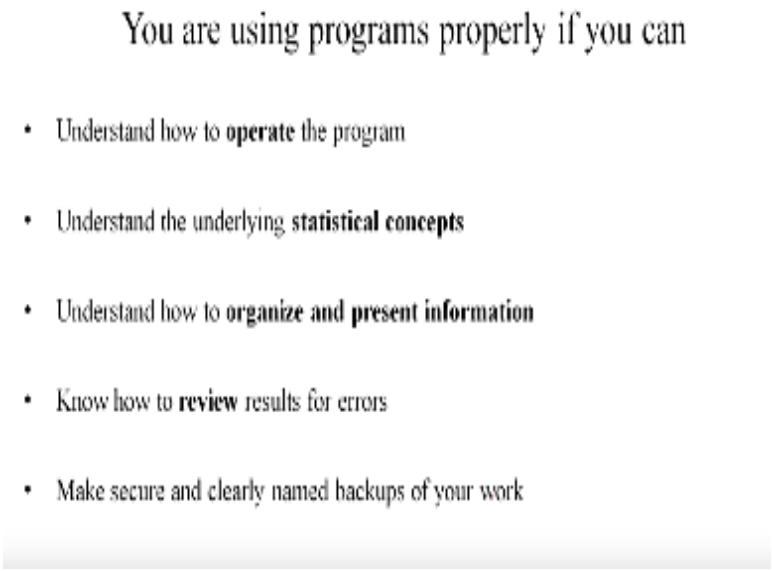
(Refer Slide Time: 27:02)



So there are different packages and tools available for preparing these charts, tables, pie charts and so on and there are some other statistical packages available. So Minitab is one of them so you have got SPSA you have got SAS you have R and so on right? Of course Microsoft Excel is

an excellent tool which can be used for data analysis. Both these use worksheet to store data of course many of you would have worked on these 2.

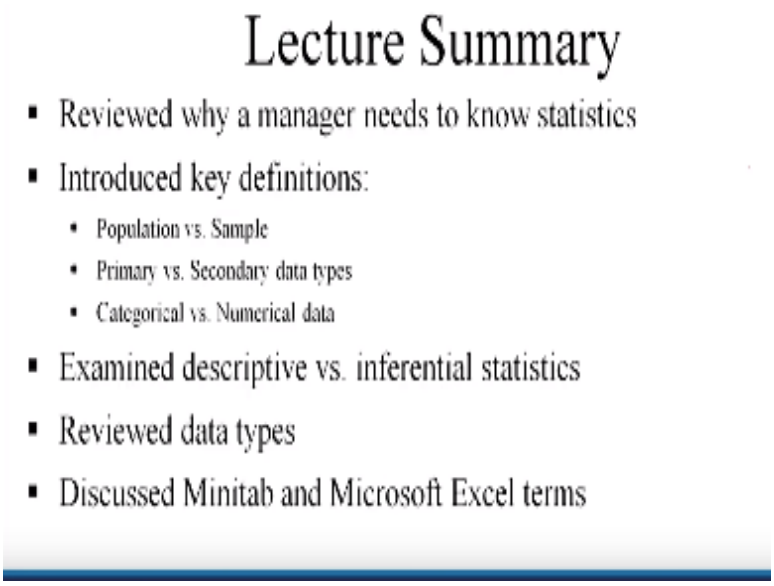
(Refer Slide Time: 27:44)



So if you are using these programs from properly then you need to understand how to operate the program. Understand underlying statistical concept this is very important you know how to operate a software. But if you do not know what is the statistical meaning of the output then it has no meaning. So you should be able to understand how to organize and present information the output which you are getting know how to review a result for errors.

Because there is possibility of having error in your results and you should report the possibility of error or in turn the probability of errors. For example, if you said that tomorrow it will rain in a city then you should say that there is a probability of rain 90% right. So there is 10% chance that it may not rain right. So let me summarize what we did in last three sessions.

(Refer Slide Time: 28:53)



We looked at population sample primary and secondary data categorical versus non-categorical data. We have seen 4 types of scales we have examined descriptive and inferential statistics. We have reviewed data types whether it is a quantitative data or qualitative data and we have also seen different sort of types of softwares which are available. Maybe as I said RR, SAS or SPSS, Minitab excel and so on.

(Refer Slide Time: 29:32)

So as far as this particular topic is concerned we will have it in next session. So with this let me complete todays session. Thank you very much.