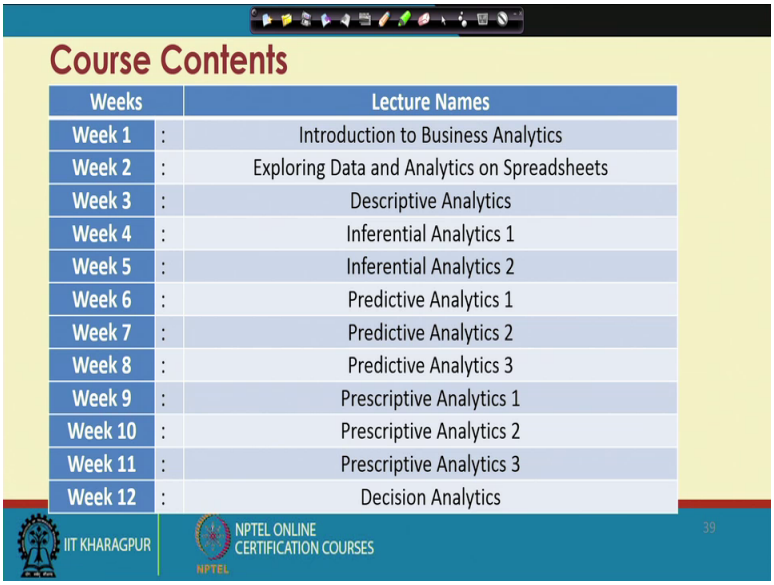
Business Analytics for Management Decision
Prof. Rudra P Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

Lecture - 03
Introduction to Business Analytics (Contd.)

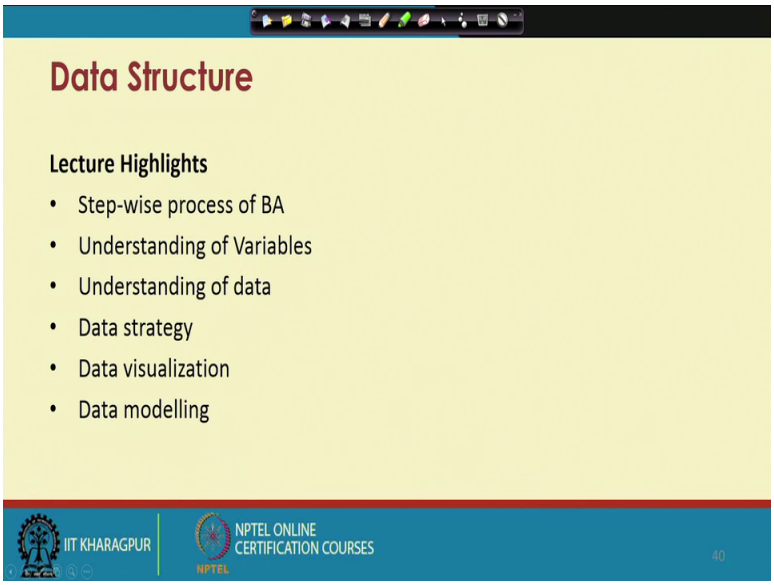
Hello everybody. This is Rudra Pradhan here, welcome you all again for this particular subject, Business Analytics for Management Decisions. We are in the process of third lecture in that too again Introduction to Business Analytics. So, today we will discuss details about our analytics requirements. So, let me highlight here is the particular structures.

(Refer Slide Time: 00:46)



So, this is how we are in the week ones and introduction to business analytics.

(Refer Slide Time: 00:52)

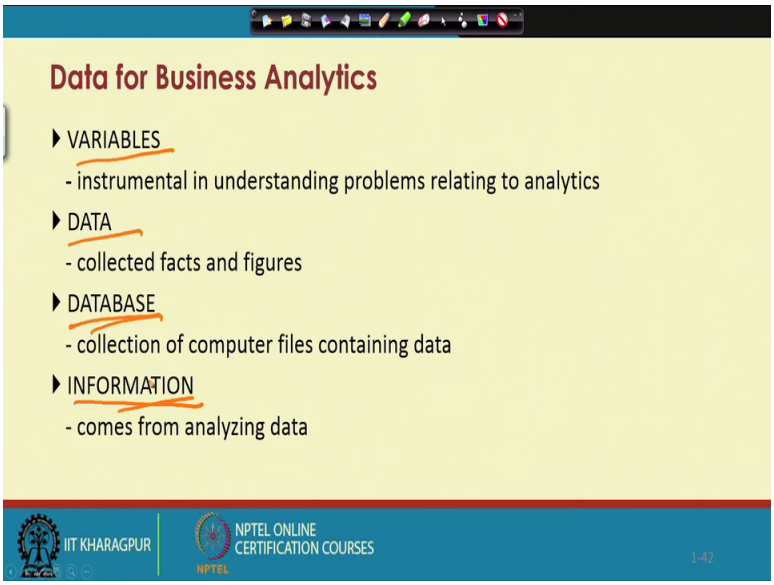


So, today's discussions we have at the following highlights. So, the first component we like to discuss is stepwise process of business analytics, then understanding of variables understanding of variables, and then understanding of data, data strategy, visualizations and data modelling. So, these are the components we are going to address.

So, the first component is the stepwise process of business analytics, and for that you are supposed to know the variables kind of requirement, then or corresponding to variables. So, here data requirement, then it will be followed by data strategy data visualization and data modellings. Let us start with the actually stepwise process of business analytics.

So, what you can go? So, the first and requirement is a so, the here will start.

(Refer Slide Time: 02:02)



So, in business analytics so, or you know one of the most important requirement or input is called as you know data. And the literary meaning of data is nothing but called as you know informations. So, informations cannot be meaningful. Until unless you specify a particular you know variable.

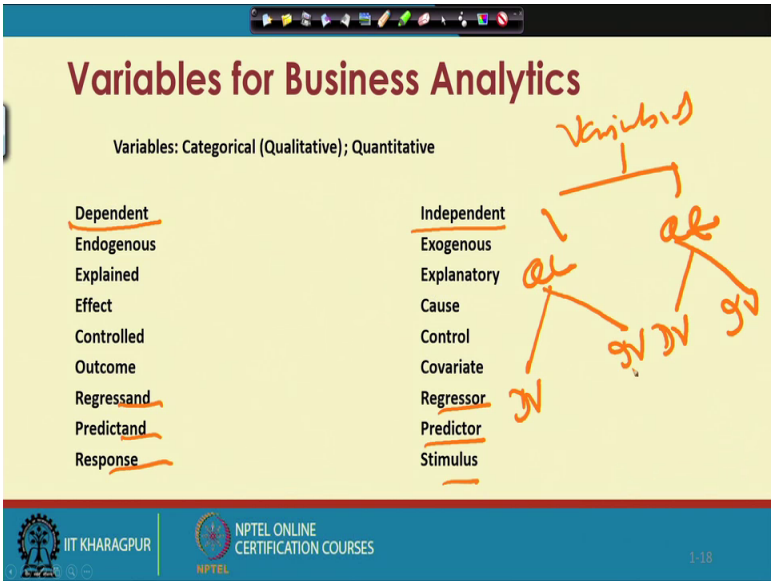
So, that is why first there is a variables indication then any information relating to that variable is nothing but called as you know data. So, that is why first you understand; what is exactly the variable concept, then data and corresponding to variables and data you have the con particular component called as you know database and informations. Where actually comes from to analyze the particular you know requirement. So, 4 items we are supposed to highlights variable data, database and informations. Database and information is not something actually very important in the stage.

So, what is important is you understand variables first, and then what is exactly the data structure. So, the variables simply called as the instrumental in understanding problems relating to analytics. So, that means so, variables and data these are the 2 components very, very tough for you know any kind of business analytics or that to any kind of business related problems, until unless you quantify a particular you know variables and that with the help of you know data. So, you are in you are not in a position to pick up any techniques, or you are not in a position to use any techniques for any kind of management decision.

So, that is why so, in any case so, you are supposed to actually recognize the problem, then transfer the problems in the form of you know identifying the variables, and then collide the information corresponding to this these variables, and after that analytics can be played very fantastic roles. These are the important you know requirements and that to you must understand all these things until unless you understand properly and data properly.

So, again you are not in a position to pick up techniques business analytics techniques for you know solving business related problem, because some of the some of the analytical tools are you know very a very specific kind of requirement. For instance, let me first highlight, then I will discuss in details.

(Refer Slide Time: 04:50)



So, here is so, before we start the process. So, let me give you a kind of snapshot here. So, what is in you know variables all together. So, variable is nothing but actually something, related to a particular you know attribute and that will be changing over the time; so the questions or time or with respect to any kind of cross-sectional units. So, before we start this particular you know structure data variables. So, let me give you some kind of clarification here.

So, when we will talk about variables so, the variables can be called as qualitative variables, can be called as you know quantitative variables. So now, when you called as a variable qualitative and variable quantitative then obviously it is on the basis of some

informations. If the information is qualitative then by default the variable is called as a qualitative variable. Then if the information is numeric then by default this variable is called as a quantitative. For instance, let us say gender is a kind of variables.

So now gender reporting's can be male and female then by default. So, it is called as you know qualitative variables. Let us say profit is a kind of variables. Where you know we are putting actually 10; 10 billion dollars or 15 billion dollars, then it is you know in numerically expressed. So, then this profit component can be called as actually quantitative variables. Now, in a business analytics is a kind of area or subject or tool basket, it can be applied any kind of environment, it can be marketing, it can be HR, it can be operation research, it can be finance, even if you know some of the engineering fields or you know social science field or the social science fields we can also apply.

As results when we apply different kind of environment; so the variables names will be you know differently represented. So, the first understanding a variables is like that the structure of quantity variables and qualitative variables; now whether it is a quantitative variable or qualitative variables. So, the second understanding or you know requirement is; so the structure of actually dependent classification and independent classification. Sometimes some of the analytic tools the need you know there is requirement of what we call as you know classification of dependent variable and independent variable.

For instance, we have multi variate tool baskets and then some of the tools can be called as you know dependents techniques structure and some of the tools can be called as interdependence technique structure. And when we apply any kind of dependence technique analytics or dependence techniques or dependent analytics, then that time there is a most that you know you too first declare which one is the dependent variable and which one is independent variable.

So, in the case of you know interdependence technique like say correlation covariance. So, we may not require actually any kind of distinction between, or you know difference between or the kind of declaration of dependent variable. And independent variable but some of the techniques it is must this a excellent example is the irrigation tools any kind of irrigation techniques, you are supposed to declare first which one is dependent variable and which one is the independent variable.

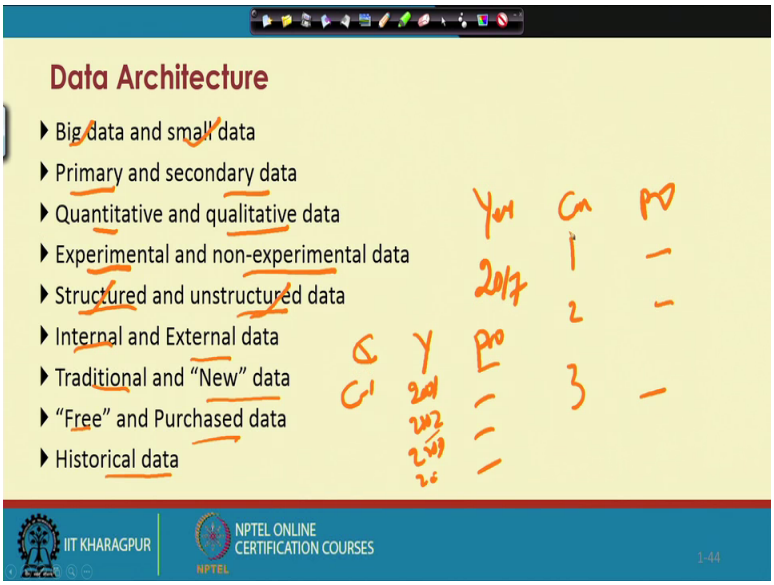
Since, we are using these techniques in various fields. So, these are you know used differently sometimes, it is called endogenous variables sometimes it is versus exogenous variables explained explanatory effect and cause then like ca control and a controlled and control outcome and covariate, then you know regress and regressor predict, and predictor a response and stimulus. So, that means, technically. So, this is actually depend and cluster and this is what the independent cluster. Technically what is actually so, the structure is like this it is a variables, and this can be divided into 2 parts, the qualitative variables and it can be quantitative variables.

Then these qualitative variables, qualitative variables you are supposed to declare you know dependent variables and independent variables. And quantitative variables you have to declare dependent variable and independent variables. Of course, it is it is all requires, when you have you know a kind of multivariate environment. So, when you when you are dealing with a single variable then the dependent classification and independent classification is not at all requires. So, when you are solving a problem where there are multiple variables are there.

So, some of the variables in some of the instances you are supposed to actually declare which one is the dependent variable and which one is the independent variable. But when you are targeting data to particular variable, then this times we are not we are not having any kind of business to you know differentiate which one is the dependent which one is the independent. The business is business of this particular classification where we like to choose a particular technique. So, for instance when we choose a particular technique say regress and then this particular division is required. When you are using a technical correlation, this particular division may not be required.

So, this is how you have to be very careful how you have address this particular you know situations.

(Refer Slide Time: 10:42)



So now, I will give some kind of hint here the question is called as you know data architectures. So, we have discuss actually first-class clarification which are given is called as you know variable. Now, any information related to variables whether it is a dependent or you know independent variable qualitative variable or quantitative variables, but that is what it is called as data. Again, we have a big understanding or you know big deal with data. So, data can be understood in following angles.

So, big data versus small data, primary data, verses secondary data, quantitative data, qualitative data. So, this is same actually when we called as a variable qualitative variable or quantitative variable it is the data which can actually a recognize this experimental data non experimental data. So, this is a experiment experimental data, and non-experimental data, then structured data unstructured data, internal data, external data, traditional data, new data, free and purchased, then finally historical data.

So, these are all actually the following requirements. For instance, this is the game between big data and small data. So, we in statistic we have a structure called as actually population and samples. So, I can put like this big data means it is a population big basket and small data means it is actually sample one say subset of a particular set. So, that means, the tendency is that if you have actually big set of or big availability of data. So, you can go for you know very in depth search or in depth kind of analysis. If you have a small data then you have a lots of you know constraint, and as a result you may

not get better insights to this particular problem. So, try to have always you know big data.

So, once you have a big data then; obviously, you can get better insights and that insights can help you lot to get some kind of better management decision. Then the second understanding must be a primary data and secondary data. Primary data means you are collecting at the root levels. You are collecting data at the root level and secondary data means somebody collected for you. For instance, any data you are collecting from a database is a classic example of secondary data. Primary data means you just go to some organization and collect the informations like; what is your name, what is your qualifications, what is your age, what is your salary like this. So, these are you know first and information you are just you know collecting. And this is the example of primary data, but you know primary data primary data has a advantage and has limitations.

Similarly, secondary data has advantage and limitations, but depending upon the situation and requirement you have to use accordingly right; then similarly quantitative data and qualitative data so, which we have already highlighted. And similarly, experimental data and non-experimental data, data can be collected from the experimental basis. And some of the data can be derived without any experimentation. Most of the engineering or science field some of the data are readily available on you know experimental basis. And we need to actually analyze. Then another kind of understanding should be structured and verses unstructured; so like this structured and unstructured. So, sometimes see if the data is very structured then you know analysis can be very fast and very accurate and you can go very your movement will be very very fast and effective. But non-unstructured data means you have to first make it structures.

For example, what is mean by actually unstructured data? That means, it may not be in a proper format right some missing data maybe available. Or some data you know not in a particular you know shape with respect to time or with respect to any kind of cross sectional units. So, that is why you know see if anywhere using secondary data, you may find that you know this is by default maybe unstructured, because that fellow might have collected for some different use. And you are you know using the data for your requirement or you know your use. So, his or her requirement may not coincide with your requirement. So, that is why you need to structure the data as per your requirement. So, that is why first you understand all this things before we go to some kind of an

analysis. Then the internal and external data this is not actually big deal this is actually when you are dealing with a kind of business environment; so a particular business. So, any data related to their organization is nothing but called as internal data.

And anything outside means anything any information outside that particular you know business is nothing called as you know external data similarly, traditional and new data. Traditional data means already available, but new data means it is newly you know it is kind of modification upgradations kind of things. And free purchase that is not important for this analytics and historical data this is very important, because we need actually the historical data for most of the business analytics tools.

Without any past informations then you are not in a position to predict actually properly to the futures. Of course, you know I will give you know better understanding after this because data can be classified into 2 time and cross sectional any data, any variables reporting with respect to time is called as a time series data any data. You know reported with respect to any kind of cross sectional units is called as you know cross sectional data.

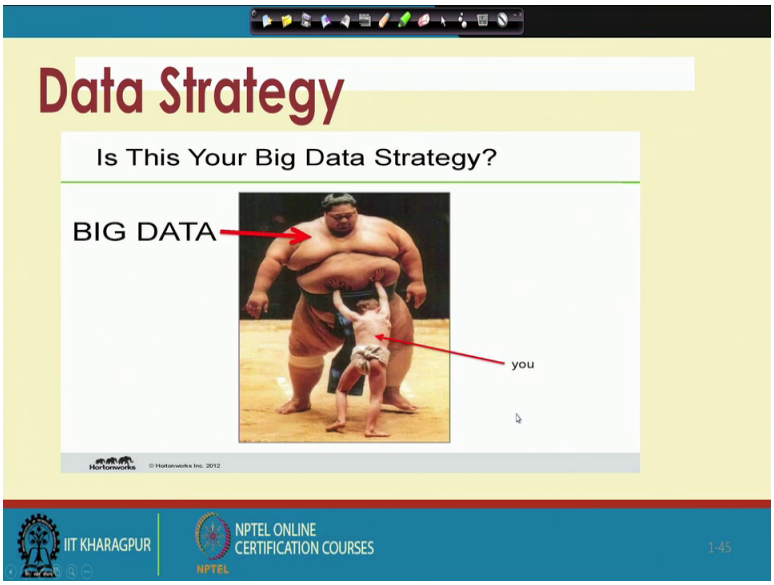
For instance, companies you know profit of you know companies for a particular years, let us say we are fixing a year 2019 or 2017, then you are you know reporting some kind of figures. You know, predicted figures or something like that. Let us say 2017. So, you are putting you know company a company b company c like that you know so for a particular year different companies like this. So, let us say this is a year, and this is the company, and this is the profits. So, putting fix fixing here year like 2017 company 1 2 3 and you know profit figures you are reporting. So, this reporting is called as a actually cross-sectional reporting.

So now, in other sides you are putting actually cross-sectional (Refer Time: 17:18) time putting year, and then you are putting actually profit here is. So, cross-sectional unit means let us say put a company one, and then company one with respect to let us say 2001 2002 2003 2004. And so, these are already reporting over the time. So, this particular reporting is called as actually times is reporting, and as a result this data is called as a time series data, and this kind of structure is called as a cross sectional structure and the data reporting is called a you know, cross sectional data. In the cross-

sectional data so, time are fixed, but cross-sectional units will be vary, but in the case of time series a particular cross-sectional unit will be fixed and time will very.

So, will discuss in details in the later stage, when we need all these you know kind of understanding or you know requirement. So, this is how you have to be structure here is.

(Refer Slide Time: 18:28)



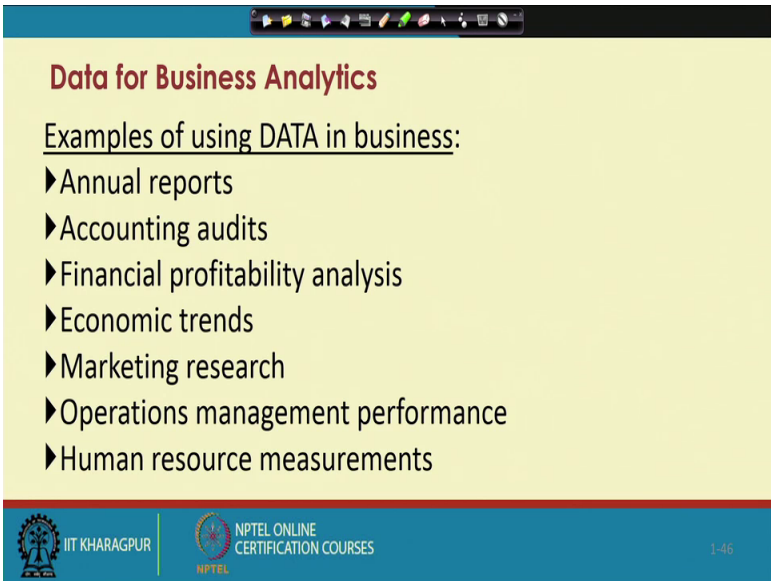
So now this is what actually strategy which I have already discussed. So, we actually need actually big data. We need actually big data that means, this stability of the analysis will be very high. The efficiency of a particular analysis will be very high. So now, if your data will be very small then some of the techniques you may not be in a position to apply. So, when you have actually more data or you know big data, any technique you can apply as per the particular requirement and of course, it should be visible one. But if you have a if you have a very small data then some of the top-class techniques as per the business requirement or you know business investigation you may not in a position to you know apply.

So, that is how that how your strategy should be always you know having you know more data, that is what is you know called as big data. So, that there is a well saying higher the sample or you know higher the data, higher is the model accuracy. Or you know higher is the tradition accuracy or higher can be you know or better you know management decision. So, that is why you must have a big data; that means, complete informations and that too in a maximum length.

So, that that is what the strategy should be, if you are you know information is lacking and information is not sufficient and not in a long interval, then the investigation may not be very effective, and you cannot get the insights properly, and you are not in a position to analyze properly, and as a result the whole process will be in a different shape all together. And that too you cannot get you know better inference, and then your management decision maybe not effective. So, that is why so, your strategy should be like that you know you must have you know high sample size or you know you must have a big data.

Once you have big data you have the kind of flexibility you can pick up the particular sample as per the requirement, you know fix the time periodic or fix the kind of fix a particular you know structure you know cross sectional structure then you know pick up these samples so it is you to decide how to take. But the availability is not there, then you have lots of you know negatives to apply this business analytics as per your business requirement.

(Refer Slide Time: 20:51)

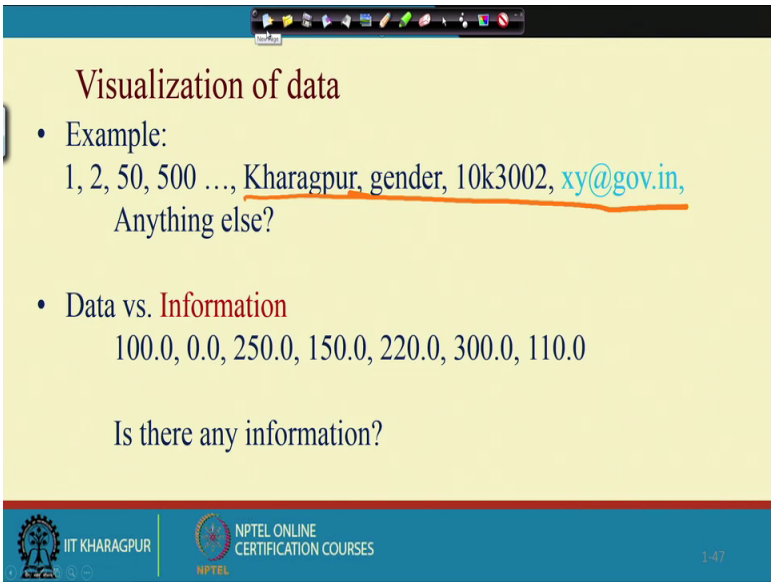


Some of the standard examples of you know data for business analytics are, you know annual reporting accounting audits, financial profitability analysis, economic trends, marketing research operations management performance and HRM's. So, these are you know classic examples through which these are all it is basically it is basically the you know data, and that too as per you know kind of requirement for you know all kind of

business activities. So, the message or you know signal is that you know, without data all these items cannot be cannot be analyzed properly, or you may not in a position to understand properly. And forget about the kind of further in-depth study by using any kind of business analytics and as per your business requirement.

So, you are supposed to actually first you know understand or you know record properly, then you have to you have to have actually proper kind of understanding. So, here is.

(Refer Slide Time: 22:07)



So, next so, this is actually the concept called as you know visualization of data. So, your data can be recorded like this. So, some of the examples are you know 1, 2, 50, 500. Kharagpur, genders, then the roll numbers the emails. So, it is a kind of standard reporting format. So, you know it is visualization of data means, what are the ways actually you capture the informations and you record the informations, and then you have to visualize and you have to properly utilize. Because you know sometimes you know let us say this particular structure 1, 2, 50, 500. This is actually numerical numeric representations right, but the other sides are you know qualitative reporting. So, that is why these are all actually what I called qualitative variables, but you know this is for instance, Kharagpur.

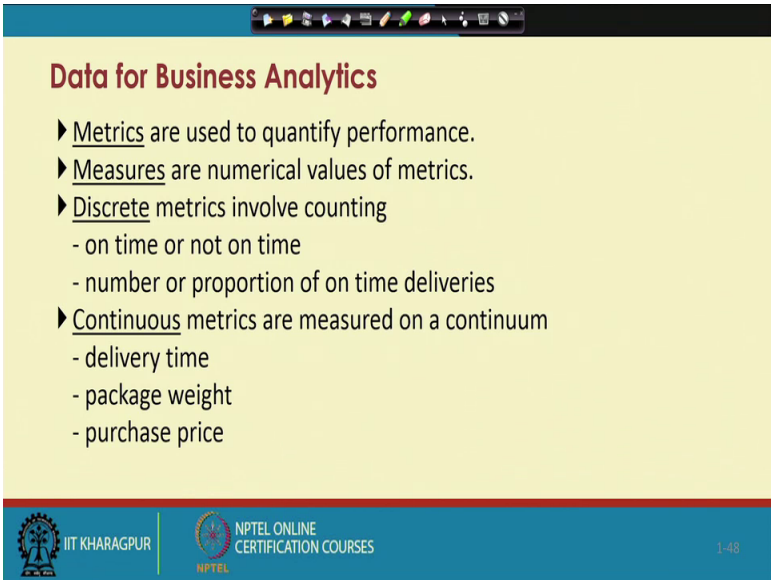
So, this is actually the variable will be placed right a gender this is actually human research component this is also a qualitative variables now this is roll number. So, this is actually qualitative information. So now, what is happening actually so, but these are all

you know data. These are all information which can actually can be used and can be applied for you know investigation and to get some kind of inference, but it is the it is not in a pro what I can called as you know in a proper structure.

So, you have to streamlined it as per your requirement, then you can apply analytics tools to get some kind of decision. For examples you know for example, let us say this is a qualitative variables and informations are you know like Kharagpur gender etcetera, but by the way when you use any kind of analytic tools or you know kind of spreadsheets.

So, these are all not numeric actually. So, some are numeric and some are notnumeric as a result you are not in a position to correlate properly. So, what is actually requirement a what is the requirement is that that you have to transfer all these qualitative information into coding. So, that means, these are the qualitative variables need to be coded first you know; that means, some quantification is required here. So, once you quantify properly then this qualitative information can be can be converted into some quantitative format, and after that you are in a position to address this you know kind of problems.

(Refer Slide Time: 24:40)



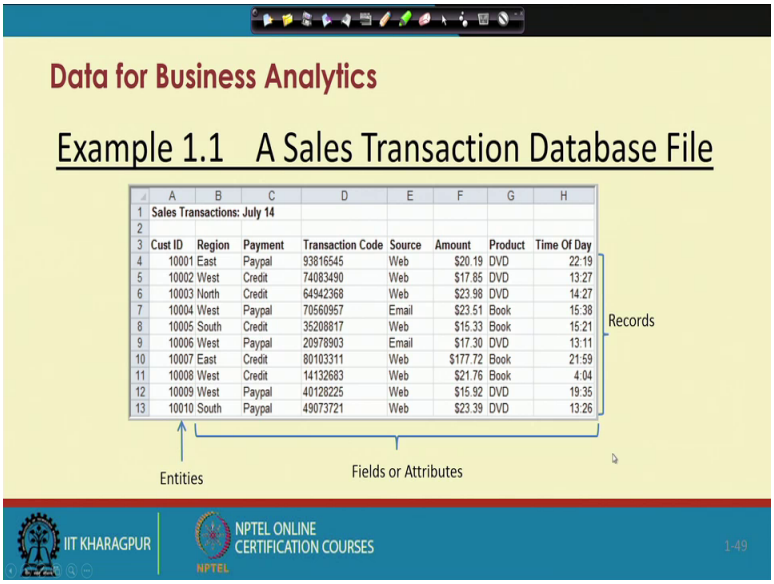
So, data for business analyticscan be can be understood from the following angles. So, it can be matrix, you know that is that is means they are you know used to quantify the particular you know performance. Measures are numerical values of matrix, and it can be discreet and it can be continuous. It can be discreet and it can be continuous. And sometimes it may be on time, and sometimes not on times. And some of the classic

examples are you know in continuous, measured on continuous calling a delivery time package weight purchase price all these things are there. So, that means, actually. So, it is question of you know metric and nonmetric then discrete and continuous.

So, you must be very careful while you know recording or you know preparing the spreadsheet or you know using this data for any kind of an analysis, because data you know the understanding of data is very important. So, if your data structure is not clear properly, then you are not in a position to analyze a you know effectively. So, you must be very careful how what are the ways actually you have recorded, and what is the particular (Refer Time: 25:52) it is actually you need actually excellent strategy how you have to streamline this particular process. Data analytics is actually very easy to pick up actually because it is a very you know very technical kind of in a thing, but this is not; so technical actually.

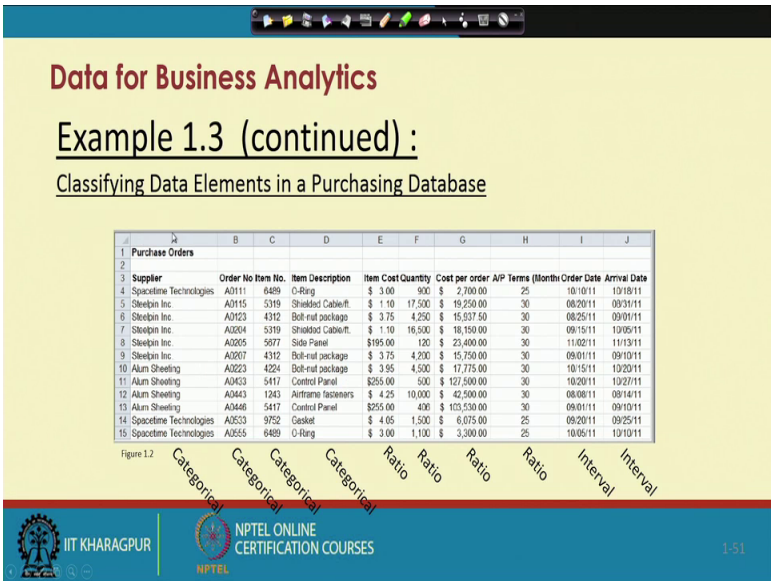
So, there is need of you know lots of you know kind of process or things to transfer into particular you know requirement.

(Refer Slide Time: 26:18)



So, this is this must be very you must be very careful. So, I am giving you here you know some standard examples. The standard examples these are all you know reporting this is a simple spreadsheet. And the data are you know recorded with customer ID within for instance this is actually customer ID.

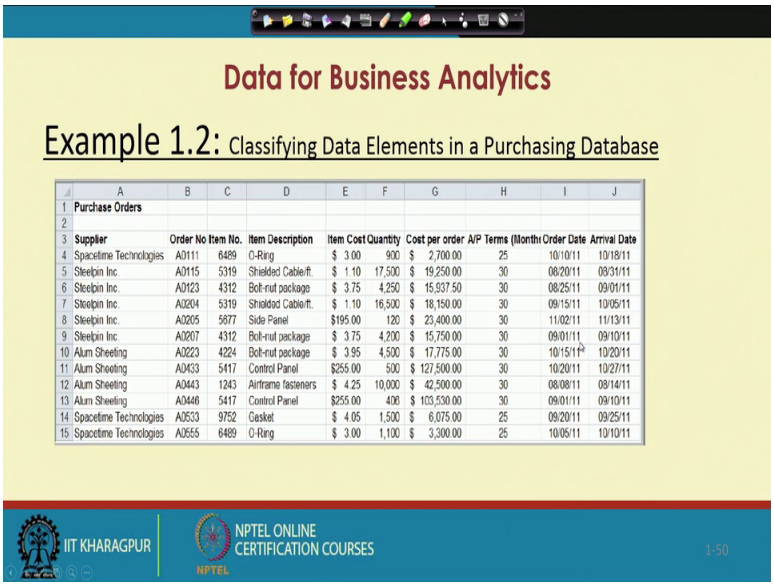
(Refer Slide Time: 26:38)



	A	B	C	D	E	F	G	H	I	J
1	Purchase Orders									
2										
3	Supplier	Order No	Item No.	Item Description	Item Cost	Quantity	Cost per order	A/P Terms (Months	Order Date	Arrival Date
4	Spacetime Technologies	A0111	6489	O-Ring	$ 3.00	900	$ 2,700.00	25	10/10/11	10/18/11
5	Steelpin Inc.	A0115	5319	Shielded Cable/ft.	$ 1.10	17,500	$ 19,250.00	30	08/20/11	09/31/11
6	Steelpin Inc.	A0123	4312	Bolt-nut package	$ 3.75	4,250	$ 15,937.50	30	08/25/11	09/01/11
7	Steelpin Inc.	A0204	5319	Shielded Cable/ft.	$ 1.10	16,500	$ 18,150.00	30	09/15/11	10/05/11
8	Steelpin Inc.	A0205	5677	Side Panel	$195.00	120	$ 23,400.00	30	11/02/11	11/13/11
9	Steelpin Inc.	A0207	4312	Bolt-nut package	$ 3.75	4,200	$ 15,750.00	30	09/01/11	09/10/11
10	Alum Sheeting	A0223	4224	Bolt-nut package	$ 3.95	4,500	$ 17,775.00	30	10/15/11	10/20/11
11	Alum Sheeting	A0433	5417	Control Panel	$255.00	500	$ 127,500.00	30	10/20/11	10/27/11
12	Alum Sheeting	A0443	1243	Airframe fasteners	$ 4.25	10,000	$ 42,500.00	30	08/08/11	08/14/11
13	Alum Sheeting	A0446	5417	Control Panel	$255.00	406	$ 103,530.00	30	09/01/11	09/10/11
14	Spacetime Technologies	A0533	9752	Gasket	$ 4.05	1,500	$ 6,075.00	25	09/20/11	09/25/11
15	Spacetime Technologies	A0555	6489	O-Ring	$ 3.00	1,100	$ 3,300.00	25	10/05/11	10/10/11

So, this is this is actually purchase orders then a anyway. So, this or we can say. So, these are all actually I am giving you 3 examples here. So, this is sales transactions and customer ID region wise then payment wise then transaction code source, then amount product. These are all you know proper recording I think you know anybody can understand you know these reporting, and this is called as you know standard reporting.

(Refer Slide Time: 27:17)



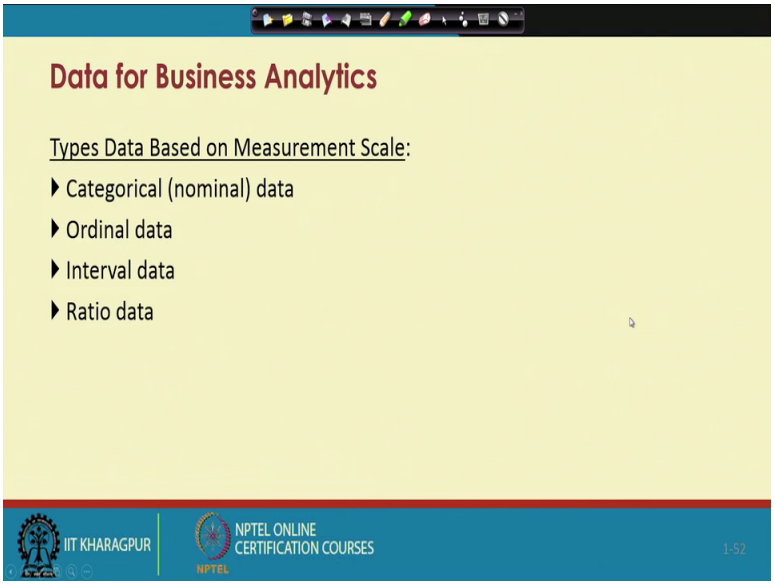
	A	B	C	D	E	F	G	H	I	J
1	Purchase Orders									
2										
3	Supplier	Order No	Item No.	Item Description	Item Cost	Quantity	Cost per order	A/P Terms (Months	Order Date	Arrival Date
4	Spacetime Technologies	A0111	6489	O-Ring	$ 3.00	900	$ 2,700.00	25	10/10/11	10/18/11
5	Steelpin Inc.	A0115	5319	Shielded Cable/ft.	$ 1.10	17,500	$ 19,250.00	30	08/20/11	09/31/11
6	Steelpin Inc.	A0123	4312	Bolt-nut package	$ 3.75	4,250	$ 15,937.50	30	08/25/11	09/01/11
7	Steelpin Inc.	A0204	5319	Shielded Cable/ft.	$ 1.10	16,500	$ 18,150.00	30	09/15/11	10/05/11
8	Steelpin Inc.	A0205	5677	Side Panel	$195.00	120	$ 23,400.00	30	11/02/11	11/13/11
9	Steelpin Inc.	A0207	4312	Bolt-nut package	$ 3.75	4,200	$ 15,750.00	30	09/01/11	09/10/11
10	Alum Sheeting	A0223	4224	Bolt-nut package	$ 3.95	4,500	$ 17,775.00	30	10/15/11	10/20/11
11	Alum Sheeting	A0433	5417	Control Panel	$255.00	500	$ 127,500.00	30	10/20/11	10/27/11
12	Alum Sheeting	A0443	1243	Airframe fasteners	$ 4.25	10,000	$ 42,500.00	30	08/08/11	08/14/11
13	Alum Sheeting	A0446	5417	Control Panel	$255.00	406	$ 103,530.00	30	09/01/11	09/10/11
14	Spacetime Technologies	A0533	9752	Gasket	$ 4.05	1,500	$ 6,075.00	25	09/20/11	09/25/11
15	Spacetime Technologies	A0555	6489	O-Ring	$ 3.00	1,100	$ 3,300.00	25	10/05/11	10/10/11

So, far as you know data is concerned. Similarly, there is another examples here supplier's informations, then order number, then you know named item descriptions,

item cost you know, quantity and you know, cost per orders like this; it is again another kind of standard reporting, but here what is the important is you know the structure of you know the kind of quantitative data and qualitative data.

So now, I will give you some kind of snapshot you know, or you know bed structure through which you can understand the qualitative data in a much better way.

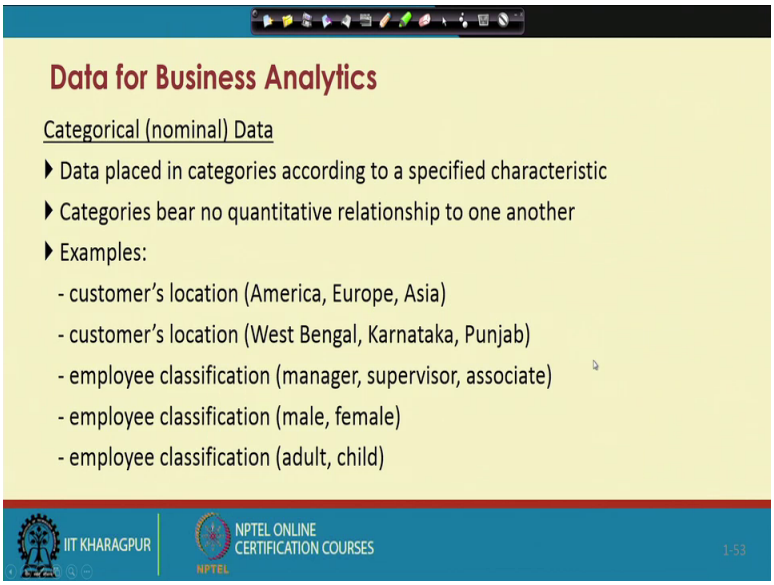
(Refer Slide Time: 27:53)



Or some kind of data structure: usually when your structure is not you know proper format, then we use a (Refer Time: 28:07) scale, and then we have to gather the informations. And accordingly a accordingly the data can be actually recorded and you know gathered. So, here is so, there are 4 different kinds of structures.

So, one particular structure is called as categorical data, that is nominal and that is exclusively actually kind of qualitative in natures like you know gender religions etcetera, and then second one is the ordinal data, third one is interval data, and forth one is the ratio data. So, these are you know 4 different strong classifications. So, for as actually data recording is concerned or data collection is concerned.

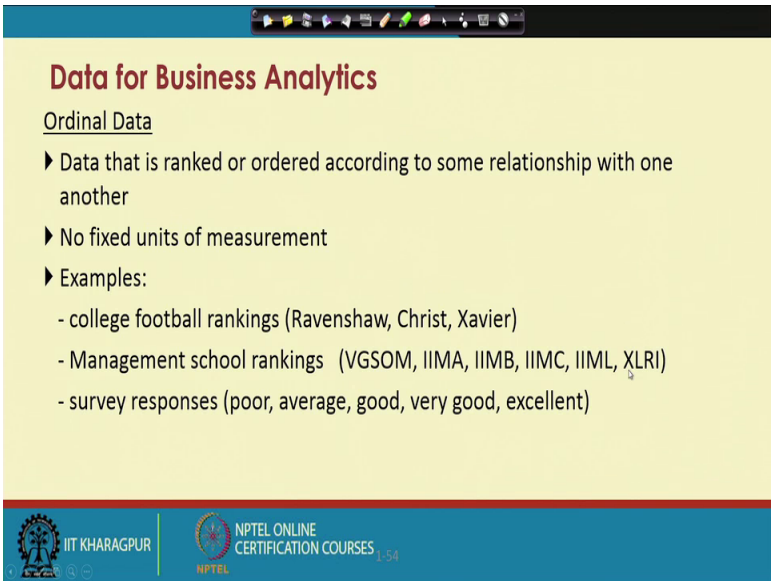
(Refer Slide Time: 28:53)



I will give you some kind of examples. So, under these 4 category. So, the first category is the categorical data that is nominal structure. And (Refer Time: 29:03) today these data are place in categories according to a specified you know feature. So, categorical datas have no quantity relationship to one another. For instance customers locations customers location with respect to different you know region of a particular country, employee classifications; then employee classifications with respect to genders. Employee classification with different kind of (Refer Time: 29:31) say adult and child.

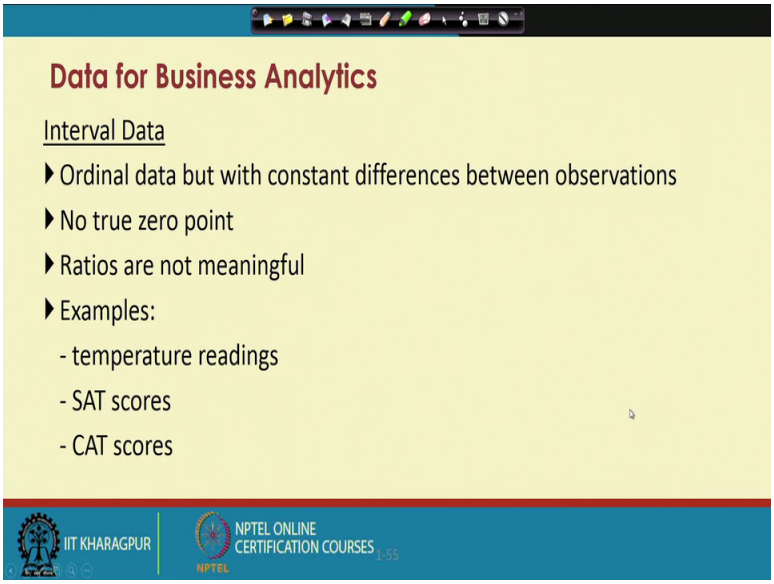
So, these are all actually the classic example of categorical data. So, while you are recording the data you must have a understanding on this.

(Refer Slide Time: 29:42)



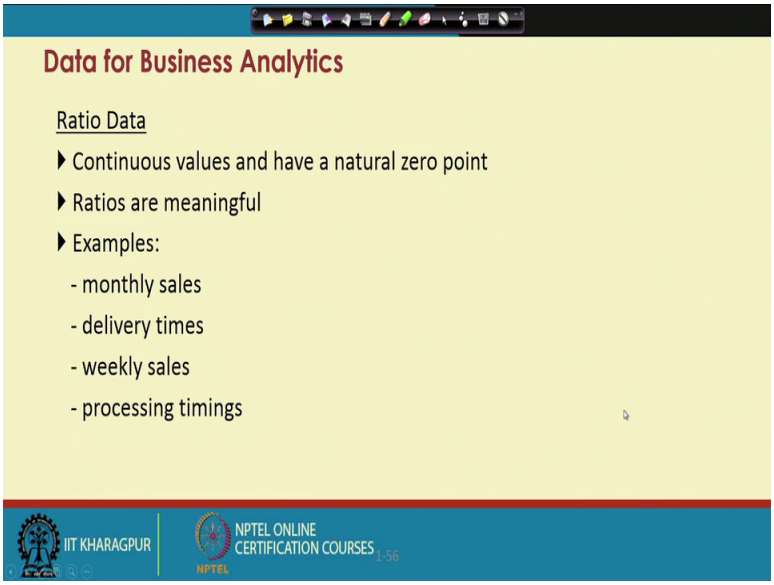
Then ordinal data that is ranked or ordered according to some relationship with one others. So now, fixed units of measurement again here; examples some of the examples are college football rankings, management schools ranking, survey responses, like you know poor average good very good excellent and so on. So, this is a kind of hierarchy kind of things you are supposed to meant. So, this is also one kind of recording, and this particular recording is it called as you know ordinal data structure.

(Refer Slide Time: 30:17)



Then, third one is called as a interval data; ordinal data, but with constant difference between observations; no true 0 point, but ratios are not actually meaningful. So, here some of the examples are you know, temperature readings SAT score CAT score, or tofel scores like these are the classic examples you can you know, you can put in the case of you know interval data. And then finally, you know ratio data.

(Refer Slide Time: 30:43)



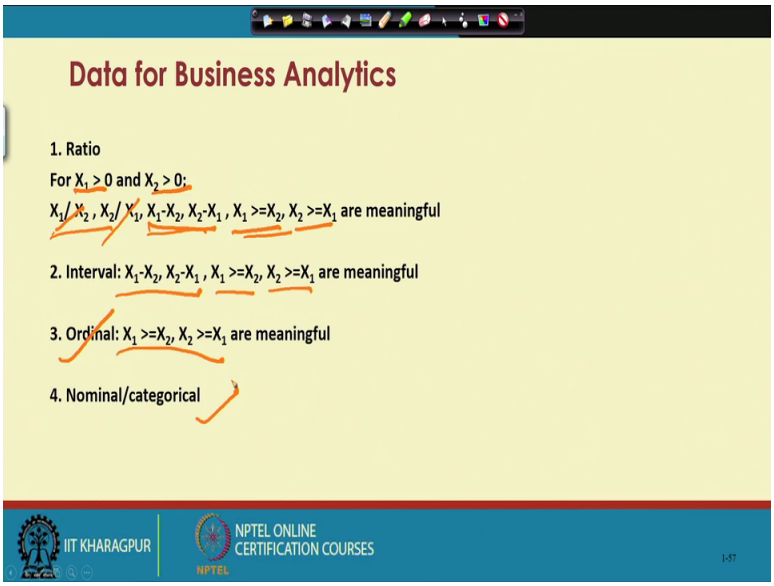
This is actually very rich data structure, and continuous values and have natural 0-point ratios are meaningful and usually the examples are you know like monthly sales delivery times, weekly sales processing times. So, these are all you know order times. So, these are all actually a the classic example of you know ratio type of data, but by the way at the end of the day. So, you must have a data, whether it is a whether it is by default in a numeric format or not in a numeric format, but the end of the day you must have a spreadsheet, having all kind of information, and if it is not in a numeric format you have to put through some kind of structure, and numerically differentiate and then that should be finally only in numbers.

So, that should be transferred in such a way that should be you know all in all must be in numbers. That is how what I mentioned actually the structure verses unstructured when you say that unstructured. Sometimes the data you know some are you know in numeric. And some are not in a numeric then you know by default you cannot just apply a

particular technique. So, when you when you choose a particular technique, or when you like to analyze these data as per your requirement.

So, be ensure that you know all these data are you know in a particular format; that means, we need actually consistency data consistency in a data recording or you know data base. If you have no consistency as per your requirement, then you may not in a position to analyze it.

(Refer Slide Time: 32:25)



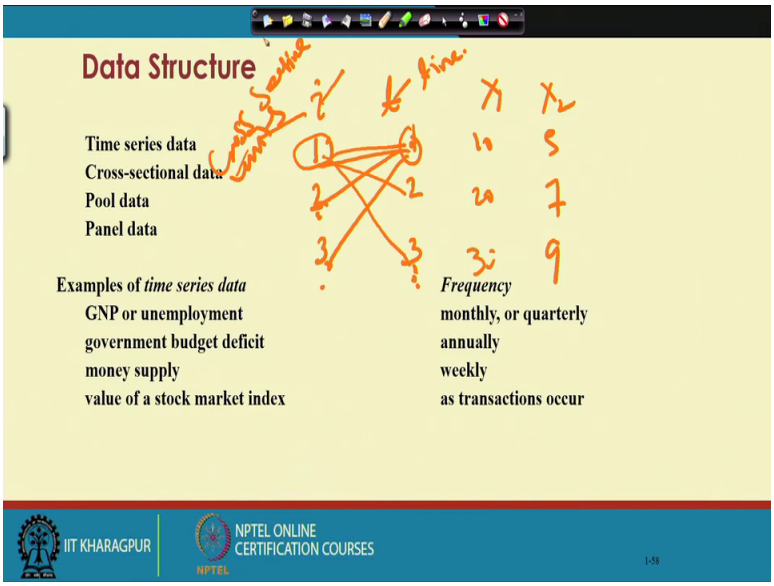
So, this is one of the most important things you are supposed to highlight the particular you know. So, mathematically there is actually you can you know on you can clarify the distinction between ratio interval and ordinal and nominal. For instance, take 2 variables to take 2 variables x 1 and x 2. And then we can have actually let us say x 1 is positive and x 2 is positive. And then you know any reporting you know x 1 or any reporting to x 2 is a kind of data.

Now, x 1 and x 2 can you can apply any kind of an operations. So, this is actually x 1 by x 2 or x 2 by x 1, and another kind of division here and this is another division. So, if all these divisions can be possible, these are all called as you know data processing. And sometimes and structured and you will make a structured by this you know process. And if any you have a data and you can process all these you know operations, then this particular data is called as you know ratio data. That is how it is called as you know rich data, data structure.

Then out of this if you now only the first one is not available, then this one x 1 minus x 2 or x 2 minus x 1 x 1 greater than x 2 or x 2 greater, than if that can be done so on. And then this particular structure is you know recognized as you know interval structure. And again finally, if the last one only actually meaningful and others to are not meaningful, and this is what is called as you know ordinal structure. And if all these 3s are all these 3 ratios and interval and ordinal are not available, then the final one is called as actually nominal or categorical.

So; that means, altogether. So, data has to be recorded in as per the requirement. And ones data will be recorded, then or you know put in a kind of structure, then you are in a position to analyze as per the particular technique and as per the particular you know business requirement. So, this is this is what actually the thing like this.

(Refer Slide Time: 34:32)



And what I have already mentions that you know, data can be actually classified into time and in across. So, sometimes you know if any variables reporting with respect to time, and that is called as a time series data in a data reporting with respect to a cross sectional called as a cross sectional data.

So now when you will be pulling time with across that is called as actually pool data. And then if you are in within the pool data, for instance let me give an example here is. This is actually I will put here i, and I will put here t. And then I will put you know variable say x 1 and x 2 like the previous highlights. So, i represent the cross-sectional

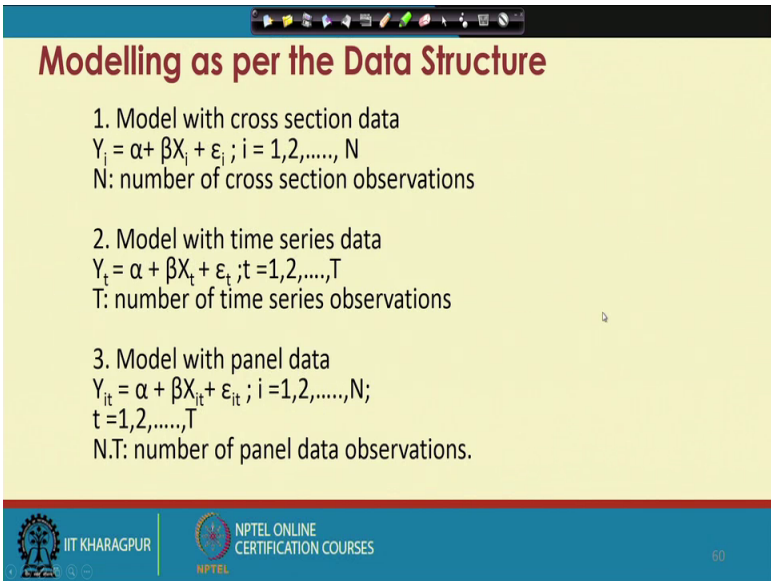
units. i represents here cross-sectional units. And then t represents here time, and then it let us say we will take 1 2 3 3 different cross-sectional units.

And we will take 1 2 3 3 different times time units, then there will be a reporting here is. So now, see here is x 1s let us say 10, 20, 30. Then I will put you now 5, 7, 9. So now, putting I remain constant. Let us say cross sectional unit means company 1; company 1 3 different time periods. So, there you know reporting the data. So now, this is what actually called as you know time series reporting. Now other way can be happened. So, suppose one will be constant here you put one constant here, then time one constant then it will be have a data like this. This will be having this will be having called as actually a cross-sectional over a reporting. So, time one particular time period, but we are reporting against you know 3 different cross-sectional unit.

And this can be also reporting so that means, it is depends upon how you are you know setting. So now, you start like this when i equal to 1. And you have a 3 different time periods, then to you have 2 cross sectional units then you have a 3 different time periods. And 3 cross sectional unit and having 3-time period; that means, technically when you are fixing i equal to 1, then you have a 3 data 3 data points. When you are fixing again you know i equal to 2, then you have 3 different times periods data. And when you are fixing i equal to 3 you have a 3 different time periods data. Then now what we can do? So, you can just copy paste copy paste copy paste like this then you know you will find finally, 9 into 9 data points. So, that particular structure is called as a actually pull data.

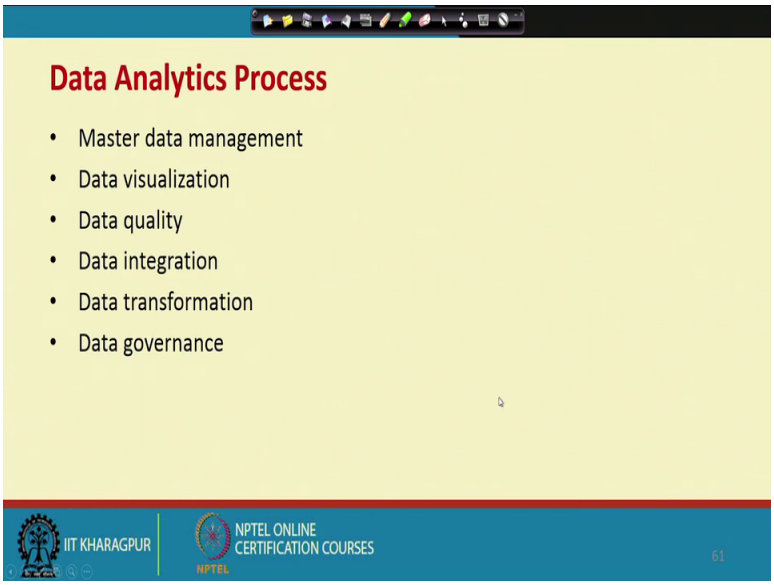
So now within the pool data when you are you know studying the impact of you know cross sectional unit and time series unit then this becomes called as actually panel data. So now so, these are the standard understanding of the data and similarly, we have a plenty of examples here. We have a plenty of examples here to know the time series data and kind of frequency. Similarly, this can be also cross-sectional data if you change if you fix a particular time period and you know changing like you know, country state etcetera like this. So, these are actually various examples of you know data with the help of all these things.

(Refer Slide Time: 37:57)



You can actually go for some kind of discussions.

(Refer Slide Time: 38:01)



So here actually some of the better understanding of the data and corresponding to the data structures we have a different kind of modelling framework, which will discuss in the later stage. And when we are connecting with a particular technique called as you know panel data modelling. So, during that times I will highlight in details about this 3 models, and then connect with you know different structures, and I will like to highlight the times how this particular you know structure and give better insight as per the

business requirement. So, these are you know final kind of understanding about the data analytic process. So, master data management, data visualizations, data quality, data integration, data transformations and data governance.

So that means, while using the data final data for your you know business requirement you have lots of things you are supposed to a address: you have to understand the data check all these you know obstacles, then make it very structured as per your requirement, then if possible you can transfer the data transpose the data. Then you can actually finally, have proper order, then you can actually in a position to apply some kind of techniques to solve the business problems as per the requirement.

With this will stop here.

Thank you very much. Have a nice day.