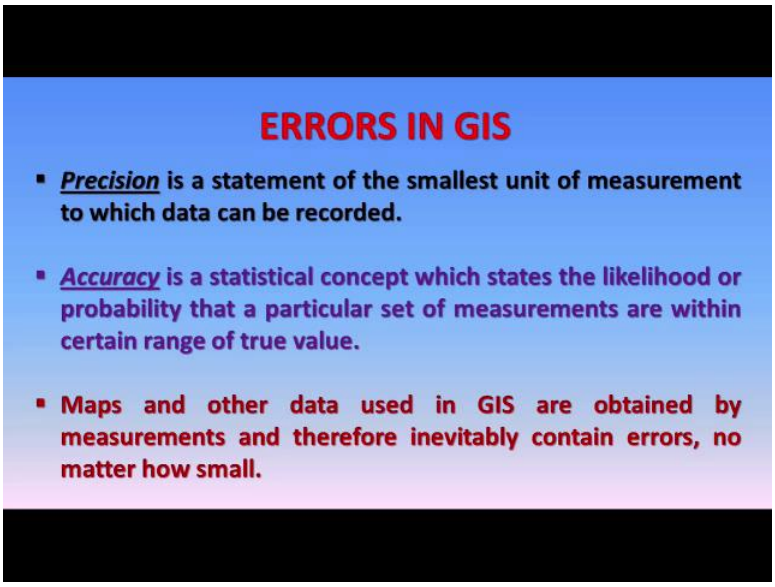
Introduction to Geographic Information Systems
Dr. Arun K Saraf
Department of Earth Sciences
Indian Institute of Technology, Roorkee

Lecture - 19
Errors in GIS and Key elements of maps

Hello everyone, I am back again and we will be discussing in this lecture errors in GIF and key elements of maps. Though these are two different topics, but I have deliberately kept in one lecture because, these are the topics which are least discussed in literature over in GIF too many people do not want to discus much about the errors whereas, this is very, very important.

That you know that error propagates in GIF and there therefore, the basic purpose of GIF not to introduce the errors to keep errors at the minimum. And therefore, we need to look into the sources of errors and how we can manage the errors. So, that we keep them at the minimum. And second is a small topic which I am going to also discus in this lecture the key elements of maps, whenever we prepare maps and we do not look towards the cartography or the elements which need to be there in almost each and every map depending on your requirements. So, these key elements we will be also discussing they are in what kind of you know these elements are and how they can be placed what is what are the importance of these elements enable.

(Refer Slide Time: 02:08)



So, all these details we will go. So, let us I start first with the errors in GIF, before that we need to understand 2 terms sometime we use them interchangeably sometimes we use them in synonymously, but these have two different the first term is precision and second term is accuracy. Accuracy is a rather easily understandable. So, I will touch that first accuracy is a statistical term or concept which is states the likelihood of probability that a particular set of measurements are within certain range of true values. And whereas the precision is a statement of a smallest unit of measurement to which data can be recorded. So, accuracy depends on user how accurately he measures using a certain tool or technique, where as precision is dependent on instrument. I will give you a example like a you might be having a one feet scale on one side your having inches another side you are having centimeters.

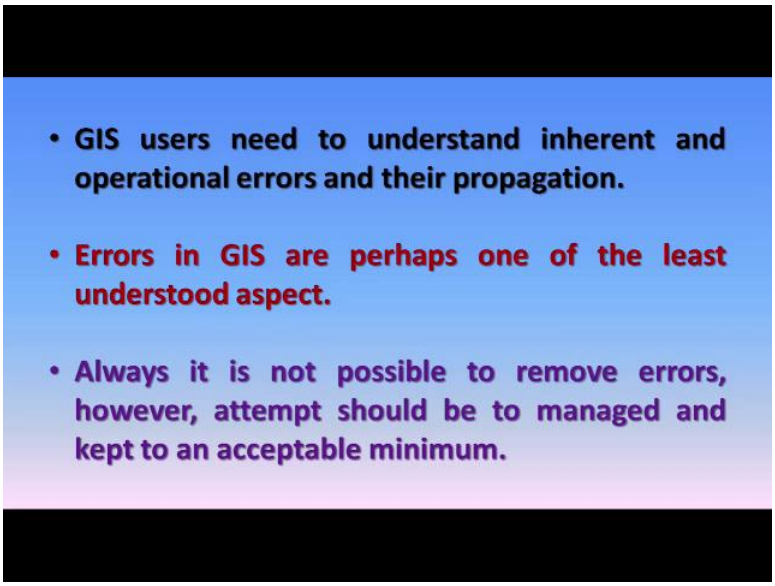
So, if you see this one inch there are 10 divisions are there and against the a corresponding one inch you would find there are 2.5 centimeters; that means, you are having 25 millimeters or 25 divisions. So, if I ask a question which scale is precise; obviously, your answer would be the centimeter scale. because that it depends on the instrument you can move a precisely measure a length using the millimeter or centimeter scale rather than inch scale because in inch scale the divisions are only 10 for the same distance that is 2.5 centimeter your having 25 divisions. So, this is what the precision is, similarly in case of like a screen resolutions or camera resolutions higher the spatial resolution more the precision you are having, and in case of digitization in case of heads up digitization we need to go for higher resolution screens to reach to that precision. Whereas, accuracy depends on the user how precisely he measures the things and as you know in simple physics experiments we were asked to measure certain things or do the experiment 5 times and then take the average of that one.

So, this is what is that the likelihood of probability that a particular set of measurements may be 5 measurements, are within the certain of true values. So, that will give you that how much accurately, but unfortunately thinks a people have been using synonymously or interchangeably which is enabled, which is not correct now certain things can be controlled like precision we cannot improve. once a data has been recorded precision cannot be improved, but accuracy probably can be improved or at the errors can be kept at minimum. So, maps in other data sets which are used in GIF are obtained by measurements and therefore, inevitably contain errors no matter how small. If you takes

a surveyed map, say topographic map those maps too are having some errors some organizations will declare that what kind horizontal and vertical errors are there some organizations do not.

But if they mention that a generally in a 50,000 scaled and topographic maps the horizontal errors are somewhere around 30 meter. So, that the object which you are seeing maybe within a 30 in reality.

(Refer Slide Time: 06:00)



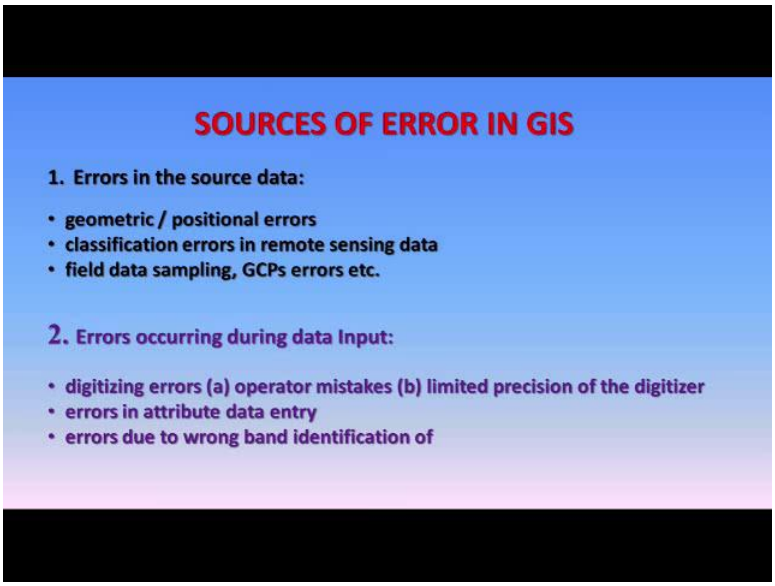
They may be within the 30 meter radius so; that means the plus minus 30 meter now as I have mentioned that GIF need users must understand the two types of basically errors the inherent errors and operational errors and their propagation. Inherent errors sometime you cannot do much if data we have bought or downloaded or purchased from somewhere or is given to us by some organization. We assume that it is very accurate, but may not be it might be having some errors. So, the better it is to have that knowledge that what kind of errors it is having. So, those are we will put them in inherent errors operational errors are committed by the GIF users.

So, those things should be kept at minimum, because you know error propagates in GIF and this is what I have also in the beginning I mentioned that errors in GIF are perhaps the one of the least understood aspect. and that that is why in literature you would in many books on GIF will not having a chapter on errors in GIF. So, that is the least understood and least discuss part of GIF always it is not possible to remove errors

because there are inherent errors everything cannot be done by a single GIF users or a one organization we have to get the data from some organization some sources. So, it is not possible to remove errors; however, attempt should be to manage and keep kept them to an acceptable minimum.

So, that we should manage in such a way that errors do not errors does not propagate one. And if they have entered in our analysis or in system or in data we should keep them at the minimum. but the information or knowledge about inherent errors and operational errors must be there with the GIF experts that you can explain to the people that this much of confidence you are having in your analysis.

(Refer Slide Time: 08:05)



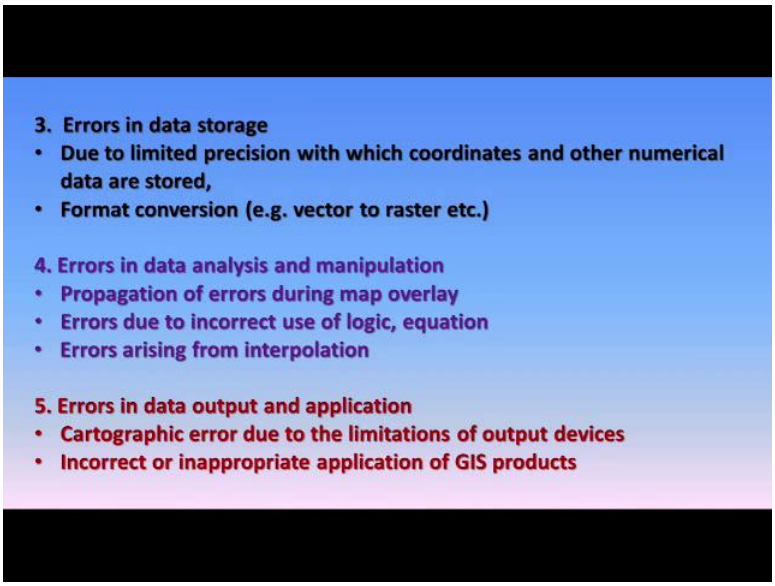
Let us look what are the sources of errors in GIF. The first error one is the errors in the source data. They might be geometrical or positional errors while doing the survey - topographic survey there might be errors in positions, and classification errors of remote sensing data some time you get a land used map prepared by some organization. If they have done the wrong classification because that land used map I have been prepared using satellite data and if a wrong classification has been performed properly it has not been checked or ground verifications have not happened. And that data if you are using that is we will put them that error in the source data. The field data sampling or ground control points ground control points, we have discussed in geo referencing part. So, if geo referencing is wrong; that means, the positional accuracy is used and therefore, these

are the inherent errors may be the sampling while the data is collected in the field, where would be the coordinate might not have been recorded properly or the measurements which you have done suppose somebody is measuring water level in the dump well.

Now, if measurements have not been done properly then that error will remain in the data for ever. So, these are the errors one has to care take care second source of error is the errors occurring during data input. Now when you know if it is dilation is done or a tributes are been thrived, at that time such errors can come and therefore, for example, a digitizing errors operator mistakes limited precision of digitizer. some earlier or then a heads on heads up digitization if the screen resolution is poor you will bring errors in the data.

So, one has to take care about those errors in attribute data entry while typing the attributes in a table if a wrong value has been typed or wrong name has been typed, that that error will remain will remain there. And we have to find out if we anytime we detect that error the best practice are check you correct it immediately, and then go to the next step and errors due to wrong band identification of remote sensing data. this is also sometimes happens because a for this number of bands are increasing maybe having say seven bands or twenty bands or 30 bands of a remote sensing sensor you are choosing a band might be wrong and I when, you make the color composites of these bands that might be wrong and then if we use such color composites for classification purposes then once the a wrong combinations have been chosen your classification is bound to be wrong and would have errors. So, this kind of errors can also occur.

(Refer Slide Time: 11:11)



Now, third type of errors sometimes and this can also occur if the errors in the data storage, due to limited precision this is one of the mistakes which people commit generally when they are a bringing the data attribute data specially. So, have to declare the properties of the your field in your data base and that property that how much precision the data will have will require when data comes, say two places after decimal or four places after decimal and if you have not kept and you have kept just a simple without any decimal places after decimal then you are truncating your data. And therefore, the precision of your data will get reduced. So, that is by due to limited precision with which coordinates and numerical data are stored and format conversion.

All a when, we discussed about raster to vector conversion or vice-versa at that time I also mentioned that these conversions transformations are not truly transparent. What does it mean here? Is that when you convert from one model to another one format to another you are introducing some errors. So, one should know that the data which he is going to use or using in his analysis how it has come? Whether it has come from raster to vector conversion or vice-versa and what kind of errors it has introduced. So, just say these conversions will always bring some errors in the data, those errors part will while converting we have also discussed in detail when we were discussing these conversions of the vector to raster and vice-versa.

Now, forth in this category of errors the errors in data analysis and manipulation, that maybe propagation of errors due to map due to map overlay, that is one source errors due to incorrect use of logic or equation. While discussing in the GIF analysis part especially in overlay and putting the syntax or you know the commands for analysis. If you a I also mentioned that, if you wrongly use the brackets the position of brackets have to be very, very carefully done, otherwise the wrong analysis will come, if you have not detected at that time then that error will come in your final outputs and your results will have a very you know very less level of confidence. So, every time I also mentioned that once you are having a complicated analysis or compound logic or compound syntax it is better to do it in say four pieces instead of just one go because after one step of calculation or analysis you can check the errors. So, your sure that at least one forth part I have checked it is fine then next part then next part rather than bringing a complicated equation or syntax and going analysis in one go.

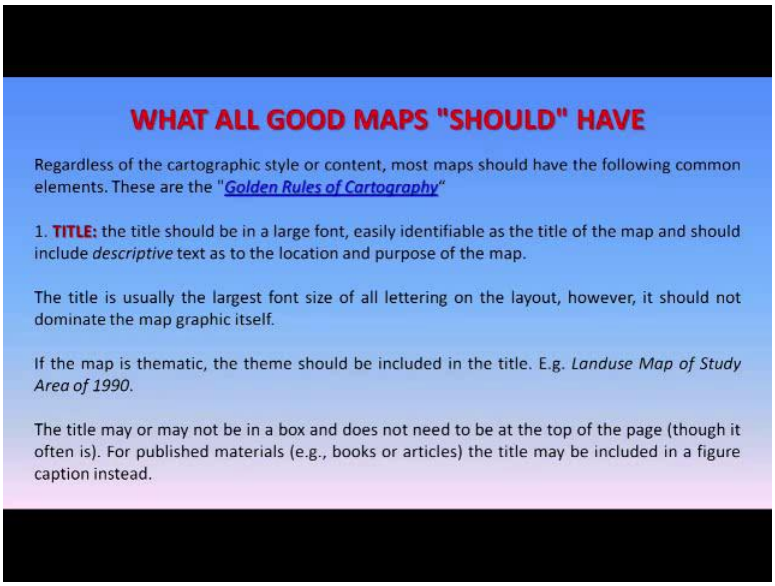
Errors arising from interpolations because in in in in interpolation we use the point data generally and they might be you know because, a different interpolation techniques will do it interpolation differently and they might bring errors. So, once you have done the interpolation there itself that would be the right time to check for the errors and, if you are satisfied then only you use that data for the later analysis if the errors are found detected immediately. Those would be corrected appropriate a interpolation technique would always help. So, once I also discussed this part when we were discussing that which interpolation technique would be suitable for my data set that can also be identified after doing certain iterations and analysis on your data using different interpolation techniques.

So, errors say in interpolation maybe due to extrapolation also errors can come and finally, in this category is the errors in data output and application for example, the cartographic errors due to which cartographic part which that the key elements of the map we will discuss in this lecture little later. So, cartographic errors due to limitations of output devices of here instead that the resolution of your output device is not as good as you should have been. So, data may be of a high resolution, but output devices are having poor resolution you are losing the quality of data might be bring in errors similarly I also discussed that when you are using is a tiff format data and raster image you convert it to the jpeg you are introducing some errors you are the quality of your

image. So, these things have to be remembered for. So, those errors remain at their minimum level.

Incorrect or inappropriate application of GIF products see GIF is not a universal tool and probably no universal tool exist in the world remote sensing is also not a universal technique. So, there are limitations of GIF which we will also discuss. So, these limitations a basically tells us that everywhere the GIF cannot be employed for a for you if a understanding is not there tools are not there algorithms are not there unnecessary utilizing GIF for that particular application will bring you wrong results and wrong name to for the GIF.

(Refer Slide Time: 17:07)



Now the next part of this lecture is a key elements of a map you know this is a generally what happens that people after analysis they have done it and they bring in map which might not be having all the key elements and then people, will start asking question that in which direction is the north what is the scale which area and so on so forth. What is the legend, what is the unit of your legend and so on so forth?

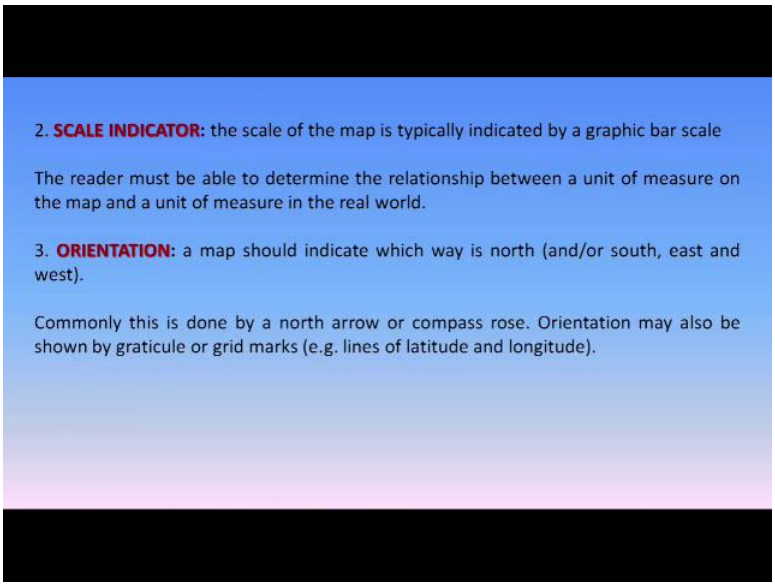
So, all these things one should keep in mind while preparing these layouts or may final outputs because you might have done very good analysis your data might be highly reliable everything you have spent. So, much time and energy and resources to create that kind of analysis, but if you do not make nice output or maps you have missed everything almost. So, it is always to you think in this direction that what are the key

elements should be there in my output map. So, that whenever it is presented to anybody either through PPT slides or printouts or anyway it should be very a complete and accurate as well. So, regardless of the whatever style one is following which is cartographic style content most maps should have the following common elements and which we put them as golden rules of cartography the title each map should always have a title if somebody is going for a publication in a scientific journal it still you can have title, but generally you carry with.

So, a is because a if a if a map is do not carry a map does not have the title then what would happen then people will ask what is this map what is what your trying to show. So, title must be there then size of this the correctors in title should be relatively quite large and it should be obvious that this is, first thing is which whichever thing is which is large the people will notice first. And therefore, one should choose the large font here and then you should tell that. For example, land use map of the study area once you have said land use map then of which area. So, title should also be a sufficient should have the sufficient information should be very brief as well as precise and if you can put a title in a box that would look also nice then a scale a scale is very, very important and that has to there and do not put a scale in ratios several softwares will support both kind of his scale, but if you put a scale say one is to 50000 and somebody changes the size of that map.

Now, that scale will not change, but what is scale the best thing is put a scale bar and you should must mention the unit's as well along with this scale bar. So, whenever a size is changed or it is projected the same scale bars accordingly will also we get enlarged or reduced and therefore, there will not be any confusion about the scale.

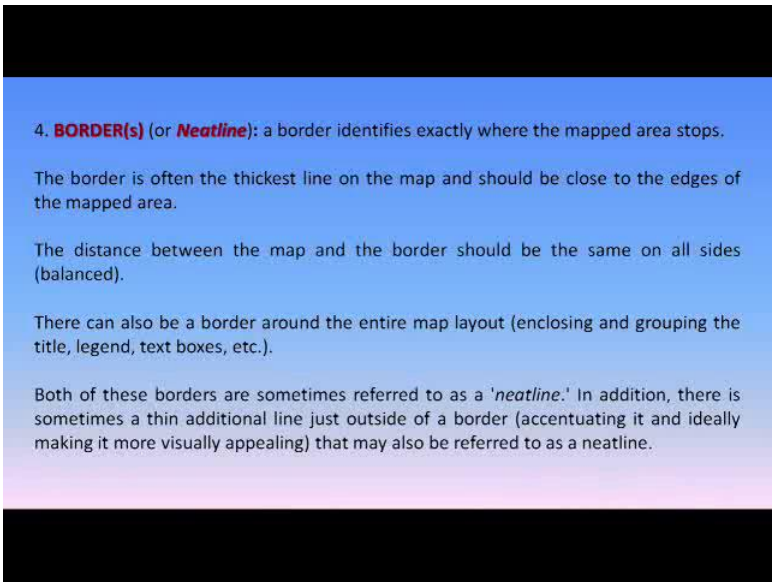
(Refer Slide Time: 20:33)



So, the scale is another very, very important key element of a of a map and as I have already said that the unit of the unit must also if it is representing a distance they may be in kilo meter meters and so on so forth. Now orientation this is depending on the requirements, whether you know maybe you as per requirement, if it is a normal a normally it has been laid out or presented then north should be mentioned. If it has been rotated then accordingly the north direction should be shown.

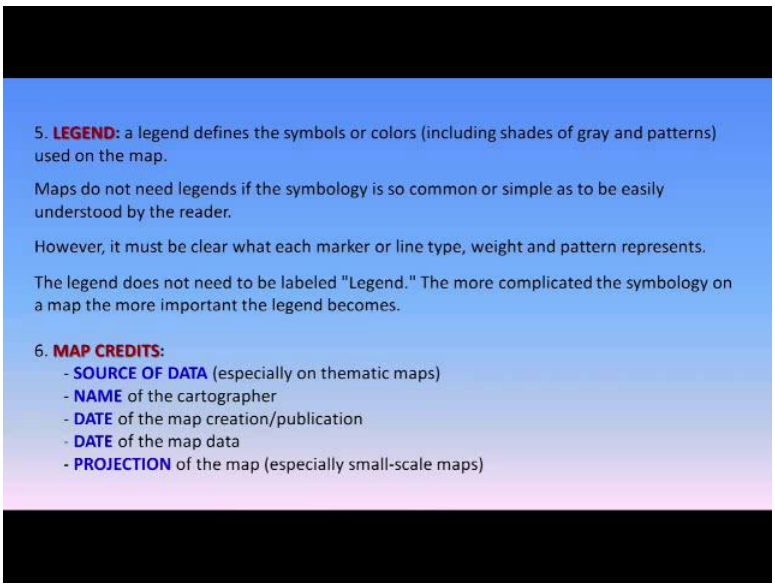
Generally people avoid putting all these symbol, but in a good software these tools are already available while going for layout these tools are available only thing you have to pull and put it at appropriate location about the north direction which is very, very important if you have rotated the map take care that you since you have rotated the map in order to fitted in a landscape.

(Refer Slide Time: 21:43)



And therefore, you must see that your north is also rotated accordingly and the forth one is the border or neat line that that also defines that this is the area of interest this is what I have worked. If you do not have any border or neat line then it looks that it is in complete kind of things that is another important thing one should have along your map outputs and then the fifth one here is the legend, this is very, very important.

(Refer Slide Time: 22:08)



Suppose you are having a digital alleviation model as a map output now if you do not have the legend how a person would know what are the range of alleviations present with

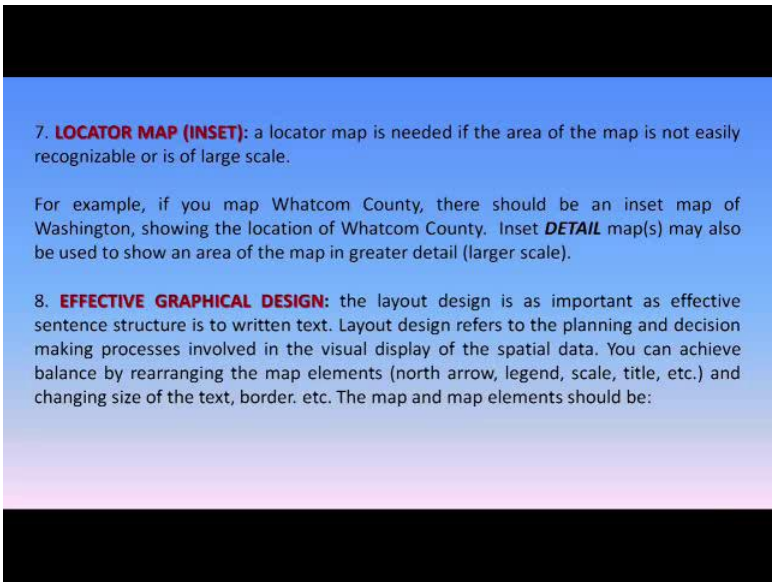
in your digital alleviation model. So, one say you are having digital alleviation model suppose range it is ranging between 0 to 1000 meter, a at the bottom whatever the colors which are representing 0 you should, that the 0 and 1000 and then unit of measurements must always be also there along with legend. So, this one has to remember that these legend has to be there it depends on what kind of maps, you are one another thing is that legend should not be very complicated very simple if the high precision is not required for the range of values do not keep you know up to three places or two places.

For example; p h value generally p h will be measured only for one place of decimal, but a because of your wrong precision declaration during due data you might have a 2, 0; 3 0s after one place of decimal. Now why it does not have any meaning? So, because of wrong precision declaration you will have that thing and it will unnecessary clutter your legend. So, legend should be a very precise very brief and full of with complete information now map credit's this is another important thing if you have got map from somewhere or use somebodies map modified accordingly. If you have prepared by yourself then that thing is not required it is understood that it is in your generation, but if I am use the maps say geological survey of India.

So, I exactly I would write somewhere in the map that the source is geological survey of India or some publications if I have use their maps modified it updated it and then I am using or presenting then I will write after that. So, the map credit's must also be given that that is also the source of data and date of data if it is remote sensing data one must mention that what was the date of the remote sensing data.

So, that it brings the completeness in your a map presentation and some other things can also come like who has prepared what was the date of preparation what was the date of the data date of the map and also the projection. If you are following some different projection, which is not normal then you must also mention that this is in the u t m projection or polyconic projection marketer projection depending what is the. So, that it is having complete information.

(Refer Slide Time: 25:05)



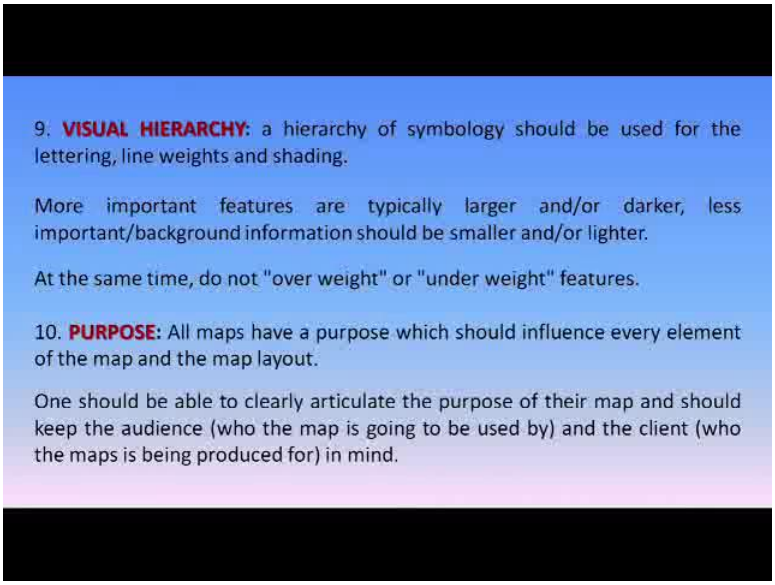
Now locator map location map is also sometimes given mainly in the publications and other places that which say map of India is there now in a small a study area which you say part of Uttarakhand. So, in you in somewhere in your bigger map you should have a inset where map of India is shown just boundary or may be boundary of different states of India and within that your study area say within that Uttarakhand maybe a square rectangular circular or whatever area.

So, this inset will give the immediate information that which is the where exactly the study area is located. So, this inset map is also very much required when I was discussing generalization here, now the generalization will be applicable because when then that map of India with this a state boundary. When you will present in a small form; that means, going for a very smaller scale you first you should go if you are having for larger scale or digitize at very high resolution now you are using for a smaller scale. So, first you go for generalization and reduce the a small details all along their boundary or these polygons and then you present otherwise what would happen that map will not look good that inset map or locator map.

So, that is another very important thing you remember and now, this you know overall this effective graphic design you know if, you are sending for publications color is not a limitation then you should chose appropriate color do not make too many if too many colors are not required do not bring too many colors, but if colors are not restricted then,

definitely you should use some colors to highlight certain things of your output map. So, that that depends on individual style in other things many extended layouts or all always available with the standard softwares which, you can which can also be used to prepare such maps the other thing is that, if you are presenting in a black and white and your using some hatching and other things.

(Refer Slide Time: 27:33)



That is also that that care must also be taken because you might be printing at a say A0 size later on this map in publication will come in A4 size.

When it is reduced choose this a hatching or patterns should not get disturbed. So, chose that pattern which will not get disturbed much, once a map is reduced or enlarged. So, that thing must be also kept in mind. Now the purpose, because overall you have to remember for what purpose, you have and prepared that map. Whom this map is going to be use if it is by a layman or a tourist then you should not have too much information on the map, but if it is for the scientific use, then for that particular purpose it should emphasis remaining things may remain in the background. So, the purpose should also be very clear should be kept very clearly in your mind, and this brings to end of this presentation.

Thank you very much.