Lecture 2 week5

Hello and welcome to this session we'll be looking at the principles of visualisation. In terms of data analysis, we have two broad classifications for visualisation. The fastest scientific visualisation and the second is information visualisation. For Scientific visualisation, typically there's no inherent geometry with respect to the data. And as a result, we can have various time durations, such as one D, two D, three d, on to and D. and the datatype could be scholar vector tensor multivariate that is a combination of scalar vector and tensor data types. As a result, this is usually associated with structure data for which continuous models can be employed to explain and analyse the data. Examples of data that could come under this category would be seismic data and medical data. For information regionalization, there's no inherent topology associated with the data. As a result, there's no structure either in terms of the time ation, this could also be 2d or 3d upon T and D. and we can use graphs and trees, scatter plots, bar charts, 3d scatter plots, and so on and so forth. to visualise such data. For Scientific visualisation, we can use line graphs, column maps is Alize, technical traces, 10 slanted sides, and so on and so forth. Examples of data that could come on the information visualisation would be named Facebook connections, the stock market, top grossing movies, and so on and so forth. Primarily we use visualisation to carry out Visual Analytics. The test, we use visualisation to understand and synthesise large amounts of multimodal data. That could be video, audio, text, images, and so on and so forth. Yeah, we have a flow diagram that gives that insight into the process of scientific visualisation. So, if we have a phenomenon, it could be a scientific phenomenon, we can create a computational model that mimics this physical process or this physical phenomenon. And based on the computational model, we have this could be a finite element model. We can run computer simulations to generate responses or Harper data for this phenomenon, and using this responses, we can have a continuous function which could be retracement to talk. And using the continuous function we can create an analytical model that could have features, objects, etc. Also from the continuous function, we can use visual mapping To create an ideal picture of what we want, which is an abstract visualisation from the abstract visualisation, which is the ideal case. We can carry out some sampling and rendering to have an actual visualisation which is going to be a district image, because we've only taken some samples from a continuous function. Once we have the descriptive agent we can do it knowledge from this image and as you can see, in this flow we have an interaction between continuous data discrete data computing computers and the human who draws the inference based on the discrete image as shown, visualisation is all about data. And to visualise data, we need to consider the dimensionality of the data is it one day is it today is the data instantaneous are time dependent that is as it changes with time such as medical data or GIS data from geographic information system such a data expectedly do have an inherent spatial domain However, there's no inherent spatial reference we also need to figure out is the data coming from a database such as chaussure databases? Once we have an idea of the dimensionality of the data we also need to consider the attributes of the data is it numeric data is a non numeric numeric data today is a natural and a integers rational numbers real numbers complex numbers are non numeric Do we have nominal values hard enough values and so on, we need to take cognizance of other characteristics as well that may be associated with the data we intend to visualise such as the domain does the data have heartburn and lower threshold certainly meets our bounds.

Once we have a good knowledge about the dimensionality, the attributes and also the characteristics of the data we want to visualise. The next curiosity we want to satisfy is Can this data be presented to portrait insight? What dimension should the presentation have? And to achieve this we are not really interested in the structure of the data. We are probably interested in the geometry and topology of the data. to present our data correctly and coherently. We need to consider following visualisation goals. Exploration we need to explore the data particularly when there is no a priori knowledge about the data, we also need to analyse the data if some sort of hypothesis exists about the data, then we can present the data when the characteristics have been well established based on the exploration and analysis. Looking at the topology of the data, again, is good to bear in mind it's the topology primarily indicates are the data points within the given data we have a distributed we are taking into communism's the connectivity between the data points and usually no connectivity even exists. When we're looking at the topology of the data, and this is generally referred to as scattered data, and because there's no inherent geometry, or topology, this comes on the information visualisation. If this is not a case, then the data is composed of cells which are bounded by grid lines. quite similar to topology, the geometry of the data specifies the position of the data. The IRS topology specifies the connectivity of the data, this can be better understood when we consider geometric shapes that remain unchanged, even when under distortion, such as we have here. So, when the geometry is stretched, squeezed and there is no tearing of the beats and the data are not stuck together. If they have been previously separated, then we can say such data possesses topological equivalence data cannot be explained by their structure as well, particularly if we have structure our data. And to do this we need to make use of special fields defined by this mathematical relation or function, where d is the dimension of the domain plus one indicates that the data varies over time. And he is determination of the range. Once we have a very vivid high idea of what his passion is, for the structure our data, then we can sample this data. by sampling, we are just taking a small part or portion of the data, which can deplete our show the quality, style, or nature of the whole data. And technically speaking, a sample is a topple, which is a point in space. And we can take as many samples as we can handle computationally from the data we have. Now once we've sampled the data, and we have our sampling data, we need to understand the distribution of the sampling data, the distribution could be regularly in our dimensions, it could be regularly in each dimension, it could be regularly in certain computation or dimensions. Or it could be erect irregular, I'll have a for all intents and purposes, we are going to assume a regular cubic grid for now. Once we have our sample data, we also need to consider our best to organise this sample data we need to ask questions such as what is the optimal data organisation? Is there a particular typological description for this sample data? What is the best way to navigate through the data or to explore this data is the data of structured or unstructured? These are some important questions. The optimality of the data organisation has to do with the time and space of the data. Once we have a sample data, we can start to aggregate all these data appears or can be placed in the coordinate space. So looking at the Cartesian coordinates or Greek, we need to understand that for the x dimension space and the way damnation space, there's an orthogonality between them. That is there are two right angles to each other or they're perpendicular to each other. And as illustrated here. For each of these dimensions, we have unit distances and each that is equidistant grid and the geometry and topology is quite implicit. So we could have an array of arrays and for the neighbouring points, then locations can be computed. An example of this is for medical diagnoses such as we have here. The data hydrogen is a CT scanner and the regular Greetings Cartesian coordinates have been used to present this and the data extent is 10 to the power of eight in the form of the data is scalar. We could also have uniform grids which are quite similar to Cartesian grids. And the consistent consist of de quoi cells. However, the resolutions are different in at least one dimension, that is the hex time measure is not equal to the Y dimension and bortolotti quantities dimension. A vivid example of this is medical imaging data that we can see style slices. So, we could have slice images with square pixels where the dimension of x is equal to dimension of y. And we could have larger slice distance ready dimensions that is larger than two dimensional x, which is equal to the damnation of y. We could also visualise structure or data using linear grids, where we have a regular topology, I'll have a for the grid points, the space in between them is irregular and the topology remains in places such as we have there, we could also use COVID linear grids such as we have here, we had to topologies to regular but the spacing between the grid points are irregular, the topology is also increasing. But if attics positions explicitly start looking at one dimension of functions or one D functions for the zero day data, this possesses an integer domain as shown in this function. Whereas for the 1d data, this possesses a real domain as shown here, generally to one manifold into space. That is why it's a function affects one dimension affects one dimension of why and it can be assumed that x and y are uniformly spaced out However, they are only known at sample points upon to N samples, some examples would be baja line graphs shown we could also have one D multivariate data for one D functions also belonging to a real domain, but wheat higher dimensional space sites also one manifold, but this time around in n space. In other words, one dimension of x is associated with and dimensions of y and the variables are nearly harmless unrelated. And to visualise this we generate a plot. We could also use a space cough in n dimensional space. The data is in parallel lines. Thank you. Do you have any questions, comments or concerns