SEPTEMBER 8, 2017

DATA VISUALIZATION

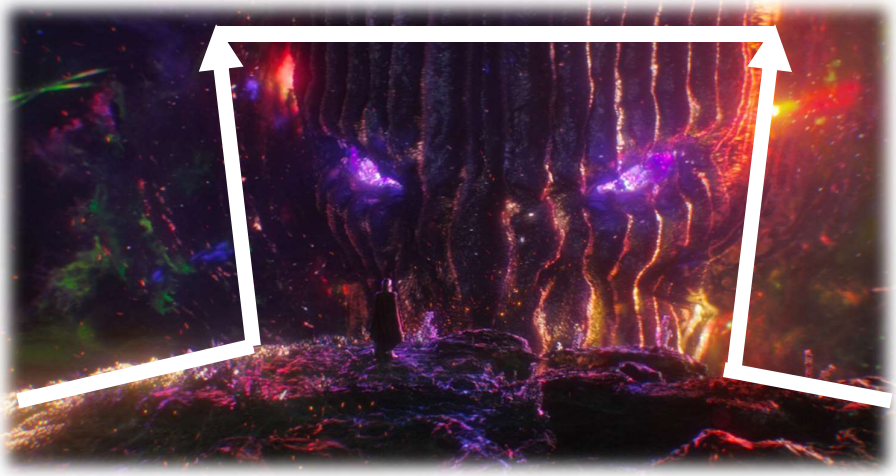
ASSIGNMENT - 1

GANDHI, ANMOL SHYAMAL
MASTER OF SCIENCE -INFORMATION MANAGEMENT
ISCHOOL

1. Dormammu – Dr Strange Visualization

- This image is from the sequence of the movie doctor strange where doctor strange comes to Dormammu and says “Dormammu I have to bargain”. This is an interesting visualisation as I feel that there a lot of factors that have contributed in final display of this picture. The complex visualisation has a lot of visual distinctions.

- I think we can divide the whole image in 2 parts for better understanding the visualisation.



- According to me the data is Quantitative as there is a clear use of X and Y axis to plot and the data points to make the distinction. Although colours are a qualitative form of data but we can still use number to represent it. They have cleverly demonstrated the use of colours to make the terrain. The use of colours indicate data points while black is just used to show no data points at that position.

- The part inside the separating lines appears to be 3d in nature and the part outside the separating lines seems like a normal background. This gives us a clue that the part inside the separation lines is formed with dense data points in comparison to the outside section. This is done to give the 3d look for Dormammu. They have used a lot of colours for distinction. The Hollow and rock shaped structure is used to form the face of Dormammu by using data points which forms a connected structure using lines. The lines are given different colour and the

hollow part between the lines is given another colour to make us believe it is 3 Dimensional. The colour distinction in the eyes makes it seem exactly like a face.

- The whole part above the background makes it seem as the face is coming out of the image. One prominent thing is how back colour is also used to show that black standing figure which is Dr Strange. So, the purpose of black here had different uses at different places. To even show no data points as well as show the main character.

- Also, there is an illusion, Seeing the size of the figure on the terrain, the terrain seems to be very big, seeing it in perspective with the figure of Dr strange but it seems to be of comparable size of Dormammu. I think it is an optical illusion due to the three-dimensional structure of the face.

- I think that the data Is processed in layers, first the background is set with the use of colours, after that that the data points are used to form the skull using different set of colours. Once the skull like structure is formed, they would have processed the terrain with colours. Overall, I think the way I have divided the image into 2 parts would be appropriate to process the data.

- This was a challenge to replicate, it is nearly Impossible to replicate this structure by hand. This is a very complex visualization which will require a lot of colours and graphic processing power to actually display that type of data.

- https://www.fxguide.com/quicktakes/bonus-luma-pictures-new-tools-for-doctor-strange/
  I cam across this link which mentions how a tool called maya was used to actually pull off this visualization by actually getting it an update called Multiverse for Maya.

2. RADIOHEAD Album Cover Visualization

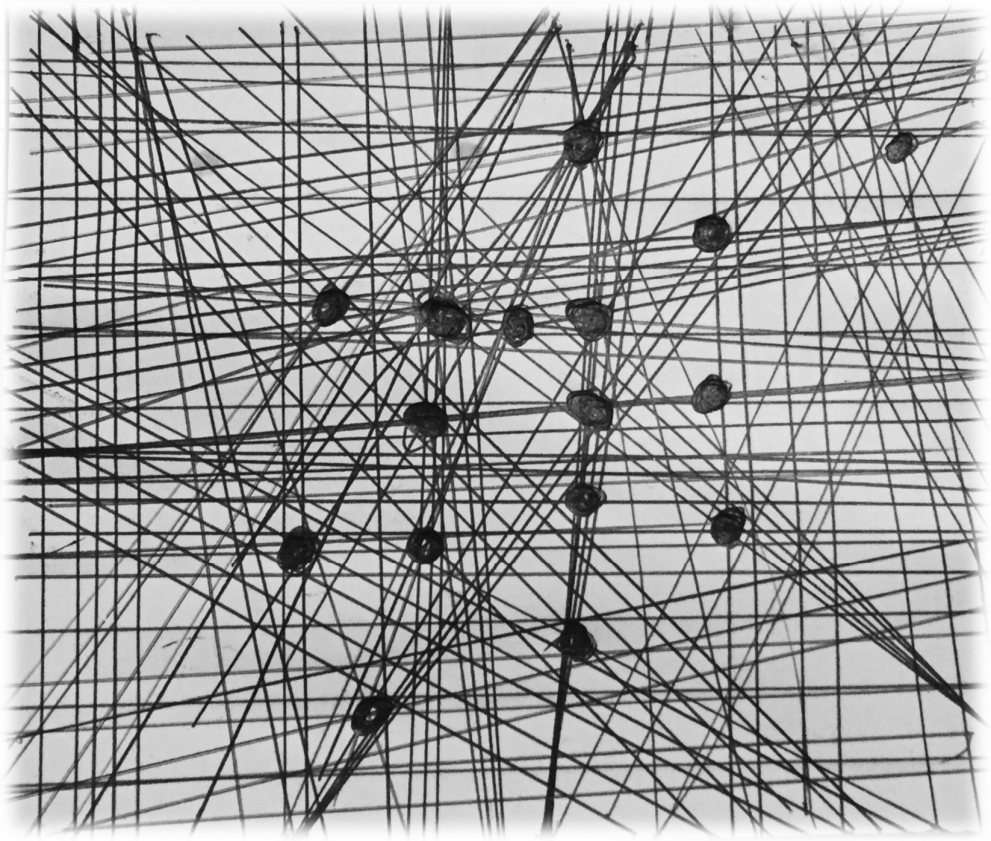
- This is the Music album cover for Radiohead rock band. This is very interesting because at first it feels like space and there are multiple lines randomly intersecting each other and forming the constellations of dense stars. The points where most of the lines intersect are quite dense and it is visually distinct. One peculiar feature about the image is that it can also be interpreted as a spider web but I am more inclined towards dense stars.

- I think here the data used Is quantitative data as the whole image consist of lines and these lines can be represented quantitatively by equation of line.

- I think processing is done in a simple way. The background is black coloured in the first step, later if we look closely this is the first quadrant with x and y axis representing points collectively. Running a graphic vector algorithm and using the equation y = mx + c they have assigned random values of m and c, positive and negative along with the points (x, y) to form these lines which randomly intersect with each other.

- As noticed in the image as the number of lines intersecting on that point increases the point starts to look denser. After large number of iterations, the whole area looks like three-dimensional, which I feel is optical illusion. It is a very complex and dense network of data points.

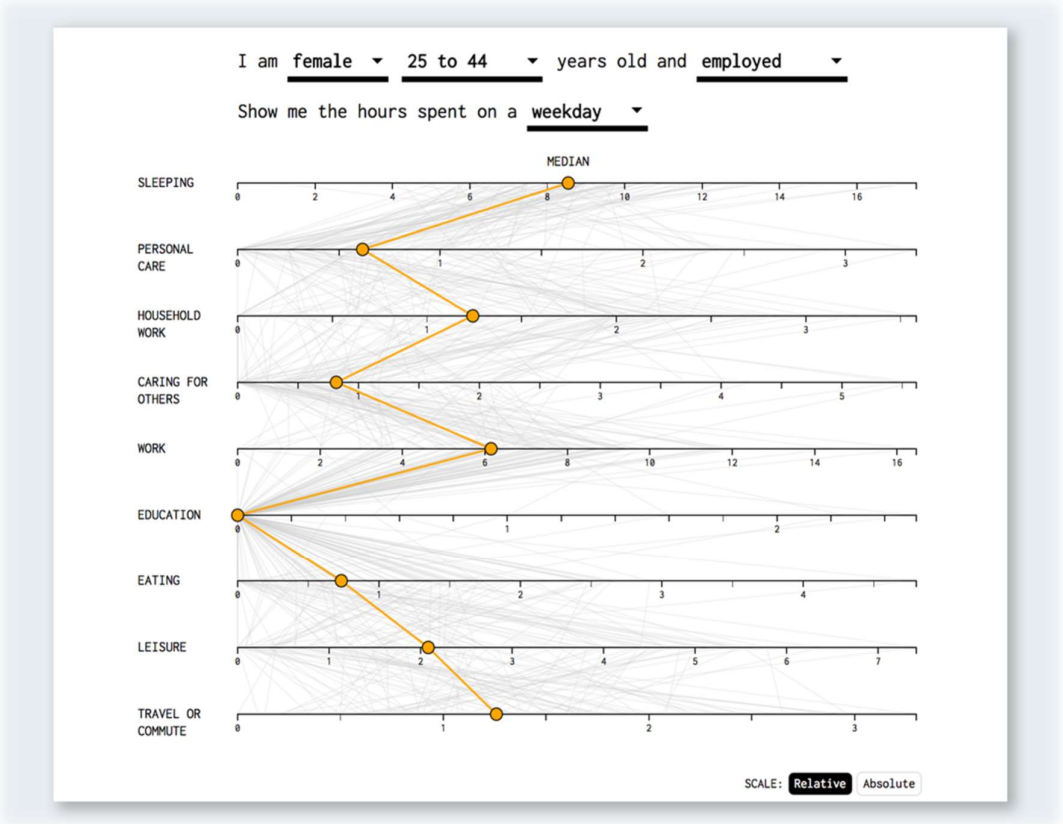
- As we can see a very interesting pattern here in this image is that, as we go near to the centre, it becomes difficult to understand what exactly our eyes visualise. We cannot look at all the dense points at the same time. The data can either be generated on fly by using a random number generator or it can also be pre-determined. Computationally generating data on the fly may be a faster process in my opinion.



- I tried to replicate the same visualization by hand but in this process ended up altering the data. To have a perfect visualisation we need to have all the information about the data points. So, in this example I tried to draw as many lines as I could by hand but drawing the same amount of lines is not possible as the data is huge.

- Also, to get the exact notion of the constellation of stars we should know how many lines we have, from where they are passing, without all that information it is highly likely that we can replicate the same visualisation

3. Daily Activities Visualization



- This is a simple visualization example of daily activities performed by female in the age group of 25-44 years who Is employed on a weekday. The visualisation has various categories such as sleeping, personal care, household etc and the measuring scale is numerical with different intervals for different categories. There are many observations displayed in the graph of which the mean time hours is distinctly displayed with the help of yellow colour, the yellow colour dot is resembling the mean of different categories and yellow lines joining the dots give us a clear visual of the hours spent in which activities.

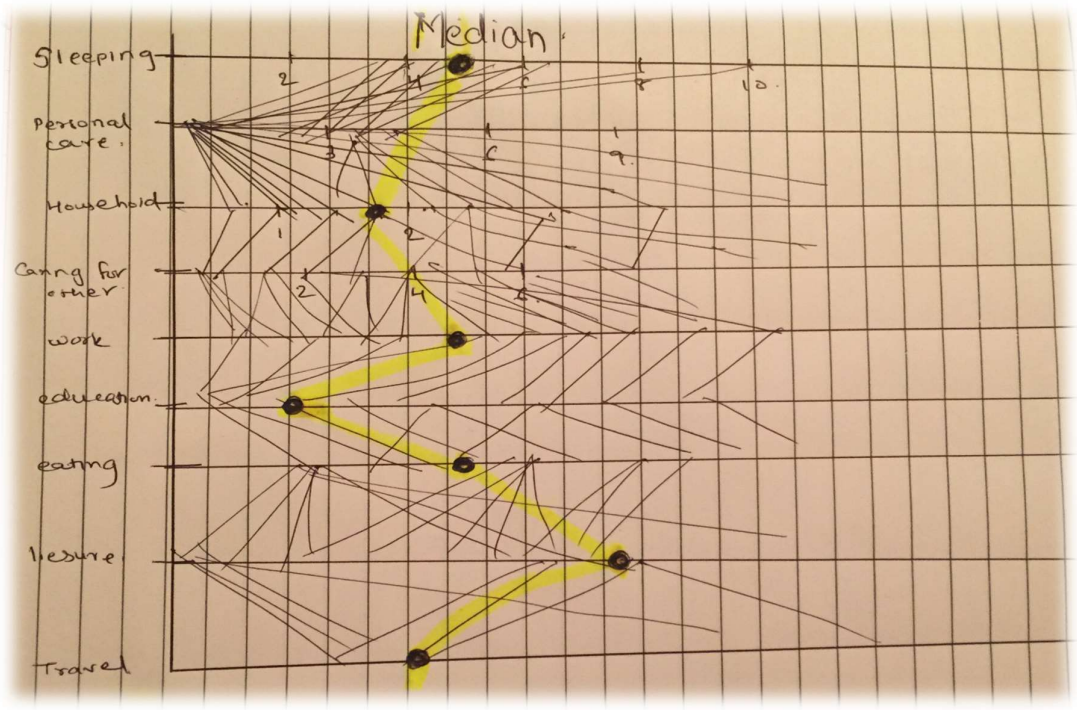
- The data I feel is taken by doing informative surveys with the women of those age group. By doing that kind of activity they must have generated a dataset of many females which may be used in this process of visualization.

- The Data here surely represents quantitative data as according to me the categories are the attributes which represent columns and the rows will be the inputs taken from the females.

- The data will be first represented in form of columns and rows after which the row data is plotted with respect to the attribute and then the dots are joined and this process continues until all the rows are processed.

- We already have the data for each category, so by using that data, we calculate the median number. A median number is calculated by sorting all the data in ascending order and then taking the middle number if the total number of observation are odd and taking the mean of middle 2 number if it is odd. Once we have the median we get the yellow dots and then we use the yellow lines to connect the dots.

- One peculiar thing to observe is that as we go away from the median the density of the plot decreases which means that most of the women spend their time similarly to the median data points. I feel this is how this data is processed.



- This at first looked very easy to replicate, but the problem here was that, I was not aware of the total number of observation used to make the graph and even while using the median I had to do an approximation to see if it could might just be the median data.

- One thing is pretty sure that if I had the data then it would be possible to replicate this type of visualisation to near perfection. To just see the original figure and assume some points to be the points of intersection would introduce some amount of manual error into the representation which ultimately changes the underlying data.