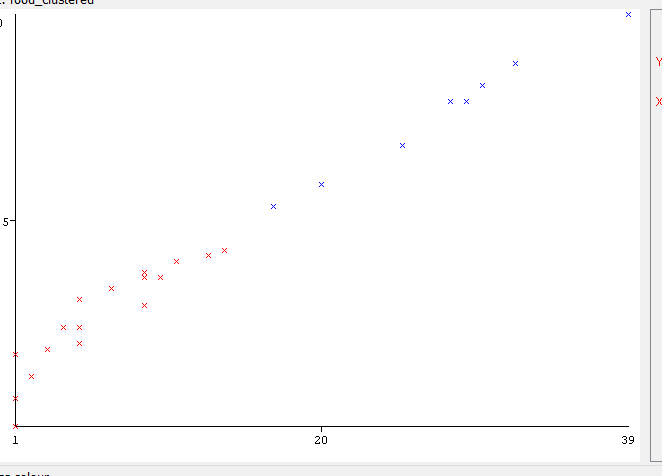
Advance Data Mining-Clustering

**By Amish Satish(Amisa178) and Fahad Hameed(fahha780)**

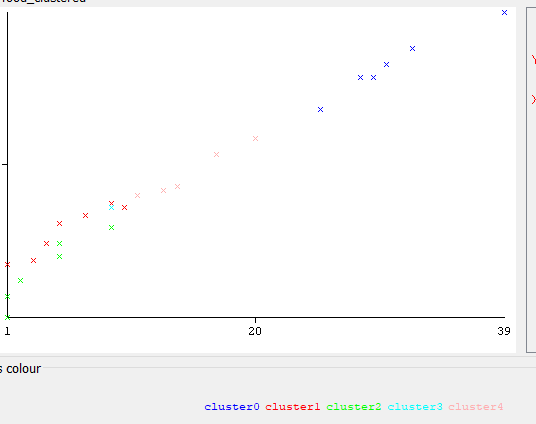
**Group 10**

# SimpleKmeans

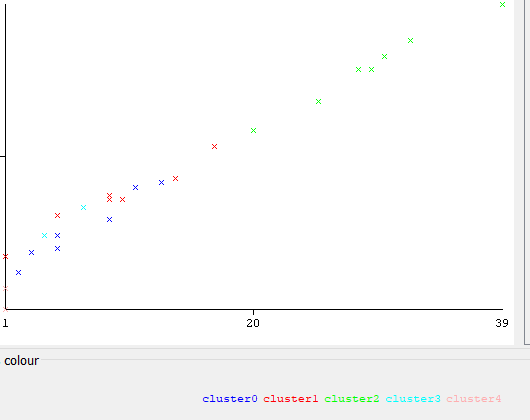
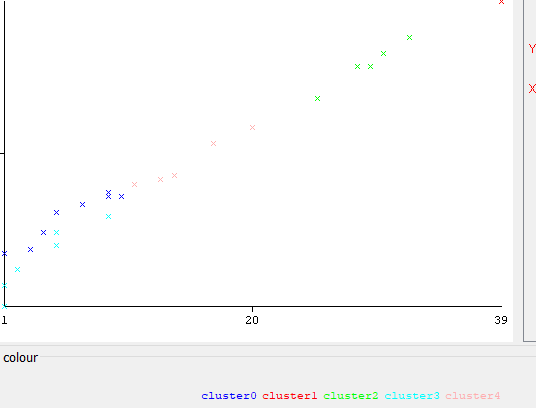
1. The “**name**” attribute must be ignored as it holds no quantitative weightage towards the analysis of the data. The following graph shows the linear relation between Fat vs Energy which seems to suggest that increase in Fat increase the amount of energy produced in the body.

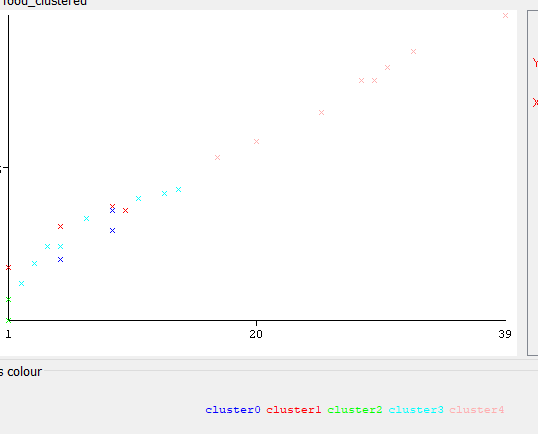


1. Different number of clusters with same seed value of 10. We get more clusters but alas is not a better representation of the data as 2 clusters was ideal for this particular scenario as it showed the best possible result with minimum ambiguity when compared to increasing the number of clusters to 5 did not bring better separation in the data points to better explain the data.



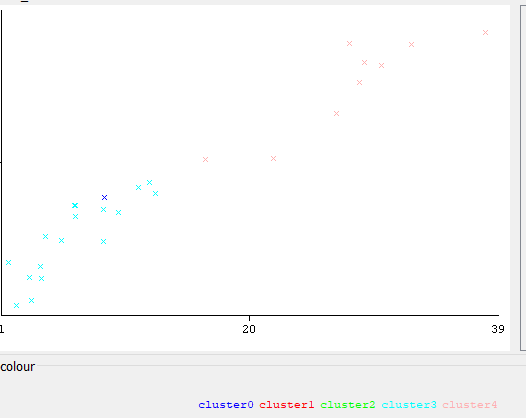
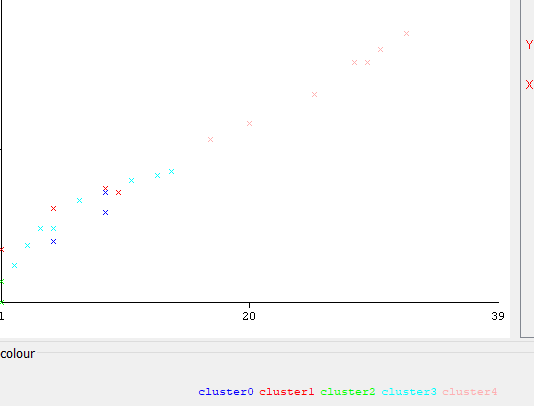
1. We used the following seed values with 5 clusters in our experiment to observe the differences between the different plots. The seed value changes the starting value of the analysis which could result in different elements falling into different clusters due to different random starting points in the data set. We can observe this in the 3 plots below with seed value at 150, 300 and 500.





1. The clusters according to the graphs are not well defined for 5 clusters as there is sufficient overlapping between the different data points and no apparent visual separation in the data plot is seen to support the categorization of the data.
2. Two clusters seem more desirable as increase cluster do not improve our results. We can name our cluster High Energy/Fat cluster and Low Energy/Fat cluster.

# MakeDensityBasedClusters



As seen above, std=0.001 and std=100 are plotted. The values show that by increasing the value of standard deviation we increase the width of the gaussian distribution which encompasses the data points. This is the reason for std=100 we only have two clusters visible whereas for std=0.001 we have 5 clusters.