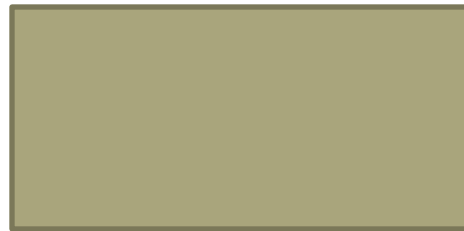


Dimensions in Clustering

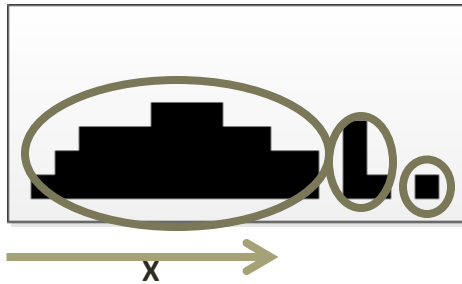
Clustering: Dimensions (1)



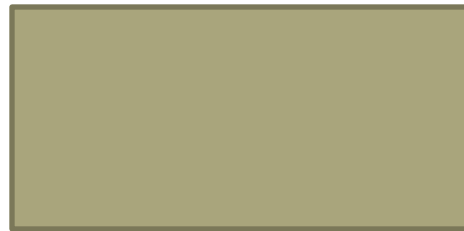
Where are the three clusters?



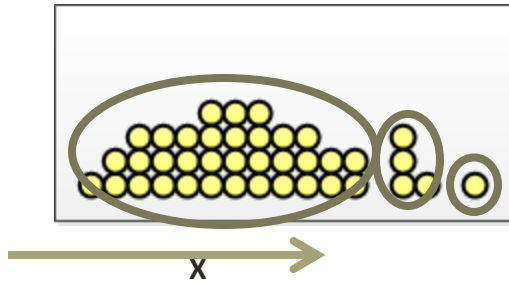
Clustering: Dimensions (2)



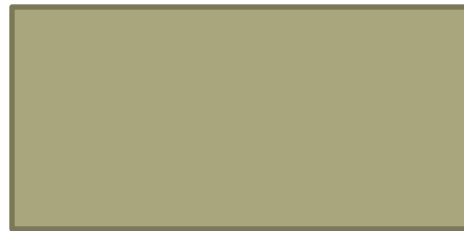
Simple assignment
based on a 1D
distribution



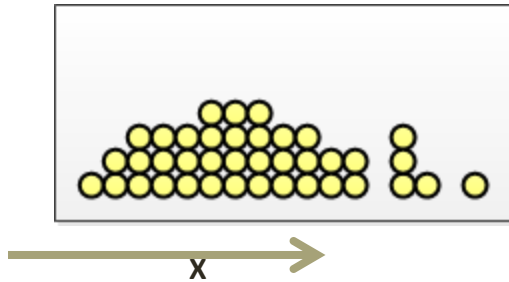
Clustering: Dimensions (3)



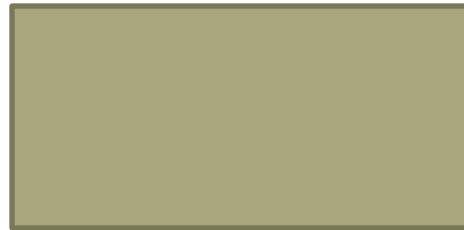
Simple assignment
based on a 1D
distribution



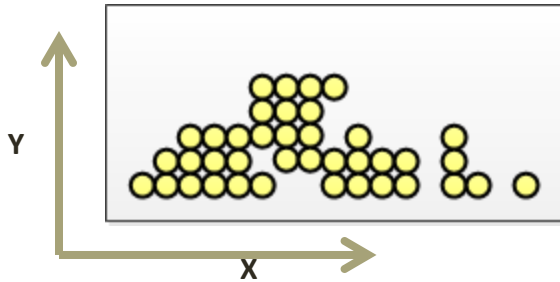
Clustering: Dimensions (4)



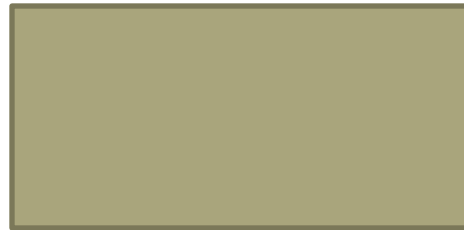
What if this was not
a 1D distribution?



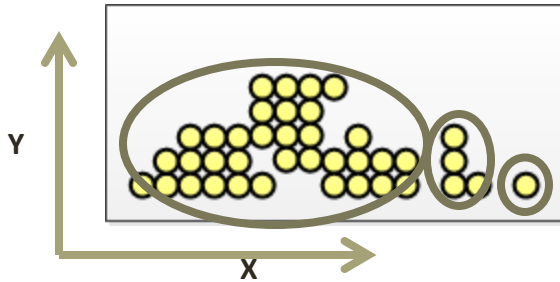
Clustering: Dimensions (5)



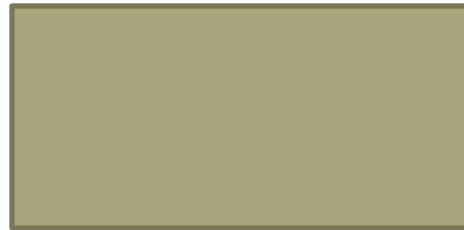
The distribution is in 2D. Some points differ in the 2nd D



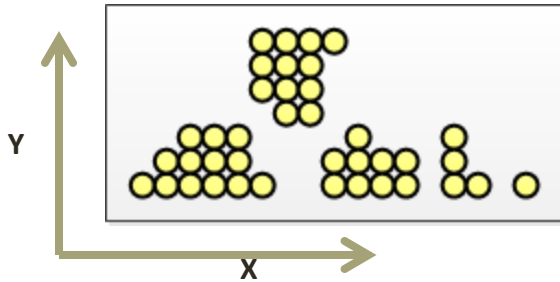
Clustering: Dimensions (6)



If the difference is minor, we still get the same clusters

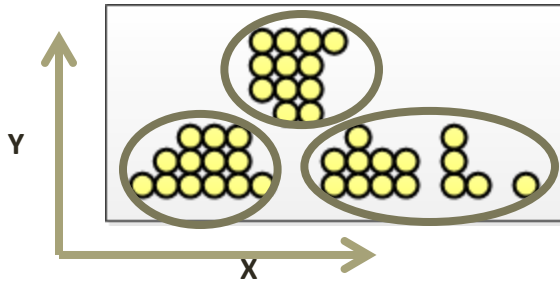


Clustering: Dimensions (7)

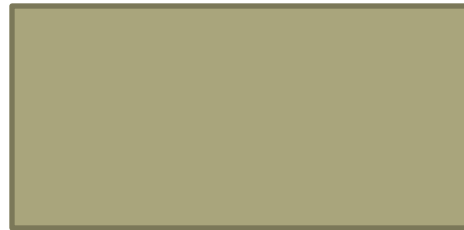


The difference could
be significant

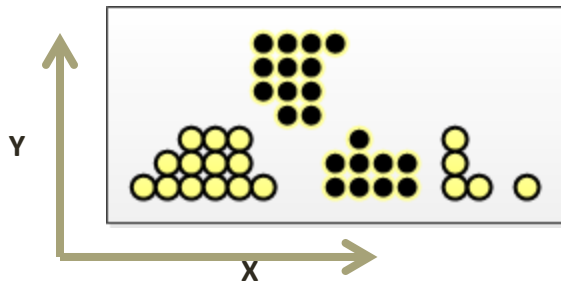
Clustering: Dimensions (8)



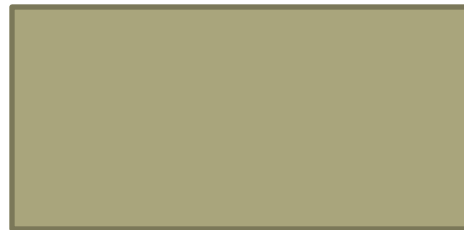
A big difference in the 2nd D can lead to different clusters



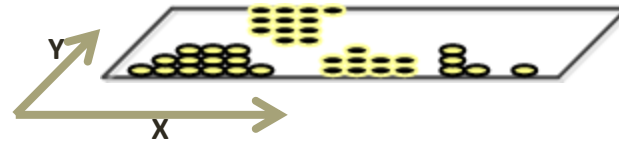
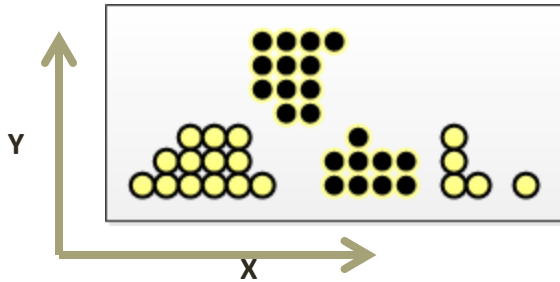
Clustering: Dimensions (9)



We can introduce another D by color coding. This is a Boolean Dimension

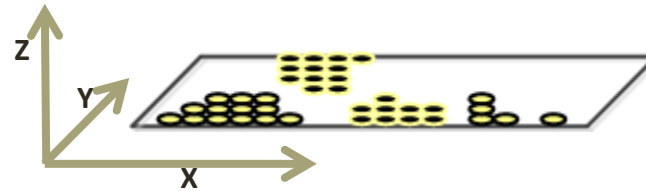
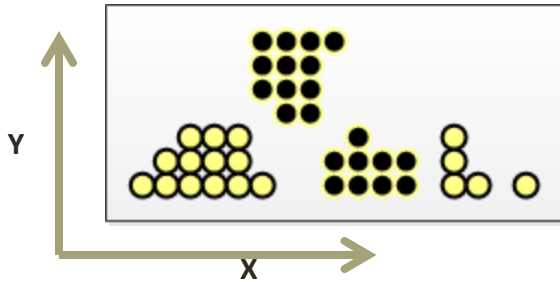


Clustering: Dimensions (10)



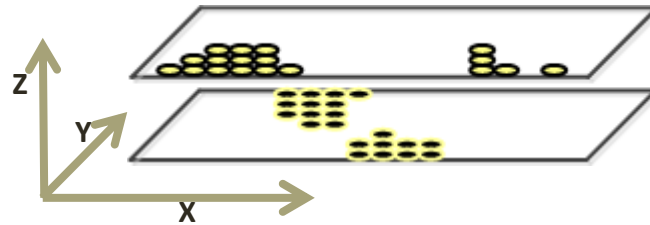
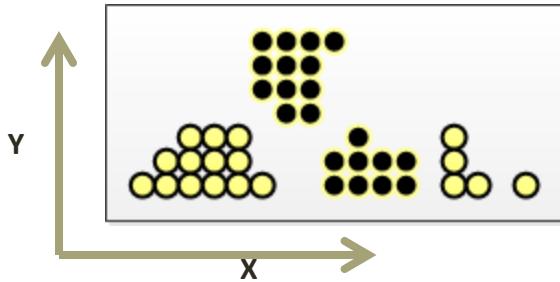
Create a 3rd
Dimension

Clustering: Dimensions (11)



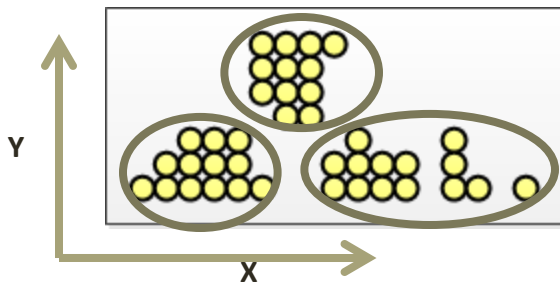
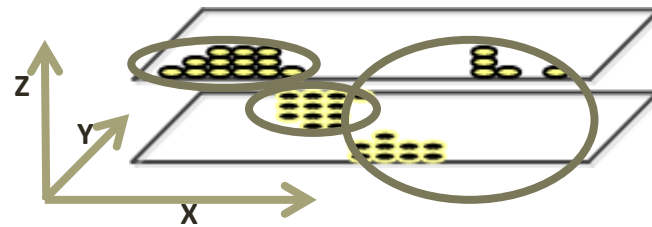
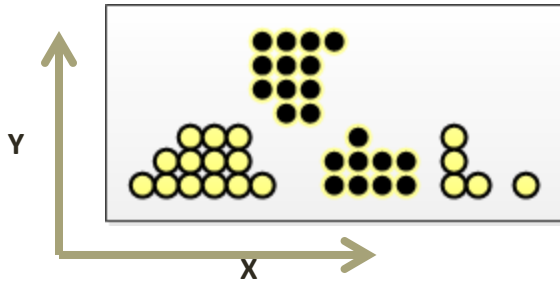
Create a 3rd
Dimension

Clustering: Dimensions (12)



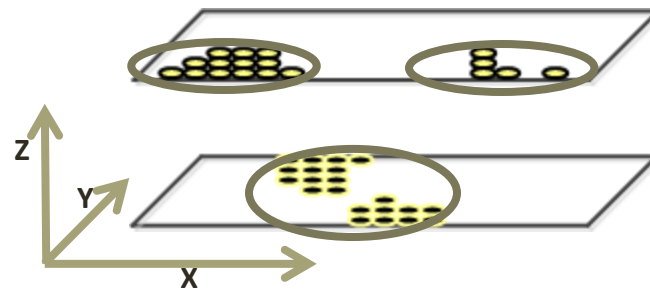
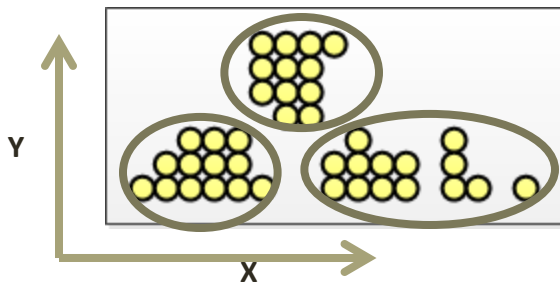
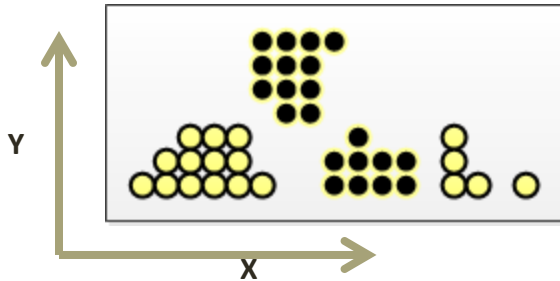
Where are the 3
clusters now?

Clustering: Dimensions (13)



If the 3rd is small,
then the clustering is
the same as in 2D

Clustering: Dimensions (14)



If the 3rd is big, then
the clustering differs
from 2D

Dimensions in Clustering