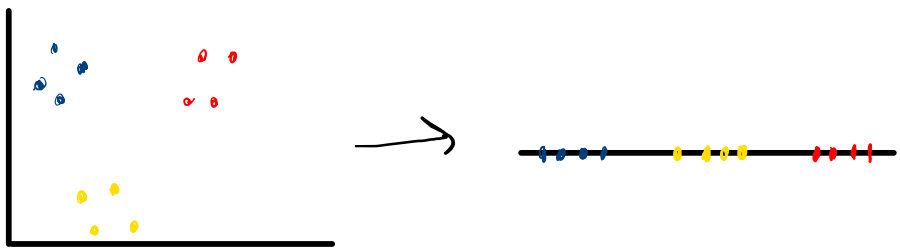
t-SNE



note projecting to an axis does NOT work

randomly place

this point attracted to red

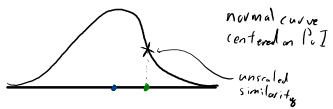
repelled by yellow + blue

physics play out, causing clustering

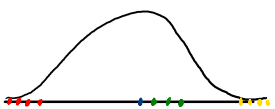
t-SNE Process

1) determine similarity between all points

Point of interest — distance d



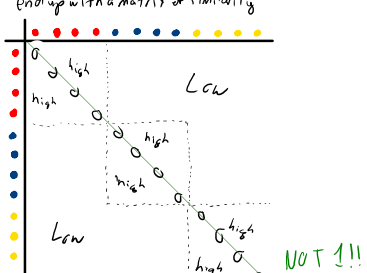
Calculate unscaled similarity between P.I & all other points



Scale unscaled similarities so they add to 1
(solves difference in cluster densities)

Calculate scaled similarity score for all pairs of points
average scores from each direction

end up with a matrix of similarity



randomly project onto number line

calculate similarity scores on line (but use t-dist instead of normal) — makes clusters easier to see

get a similarity score matrix (messy though)

move points to make matricies look the same