The t-SNE algorithm can be broken down into the several steps:

Compute pairwise similarities between data points in the high-dimensional space: The similarity is calculated using a Gaussian probability distribution. Points that are closer to each other have higher probabilities of being similar.(next slide)

To compute the pairwise similarity between two data points in the high-dimensional space using a Gaussian distribution, we first calculate the Euclidean distance between the two points. Let's say we have two data points x\_i and x\_j

where **x\_i** and **x\_j** are the high-dimensional feature vectors of data points **i** and **j**, and **sigma** is a scaling parameter that controls the width of the Gaussian distribution. The larger the value of **sigma**, the smoother the distribution, and the more likely it is for dissimilar points to be considered similar. (next slide)

We can repeat this process for all pairs of data points to compute the pairwise similarities between all pairs of data points in the high-dimensional space. These pairwise similarities are then converted into a joint probability distribution that represents the similarities between all data points. This joint probability distribution is then used in the optimization process to find a mapping between the high-dimensional space and the lower-dimensional space that preserves the pairwise similarities between the data points.(next slide)

Initialize the data points in a lower-dimensional space: The algorithm constructs a lower-dimensional space, typically 2 or 3 dimensions, and randomly initializes the data points in this space.(next slide)

Compute pairwise similarities between data points in the lower-dimensional space: The pairwise similarity between two data points in the lower-dimensional space is calculated using a Student's t-distribution.

The t-distribution is chosen because it has heavier tails than the Gaussian distribution, which makes it more robust to outliers and allows for better separation of clusters in the lower-dimensional space.

Convert pairwise similarities into a joint probability distribution in the lower-dimensional space: The pairwise similarities are converted into a joint probability distribution in the lower-dimensional space.