Imagine you're a biologist studying a new species of butterfly. You've collected data on the size, shape, and color of the wings of each butterfly, but there are hundreds of different measurements for each butterfly. How do you make sense of all this data and find patterns that could help you understand the species better? This is where t-SNE comes in.

T-SNE, also known as t-Distributed Stochastic Neighbor Embedding. It is an unsupervised, non-linear technique that helps us visualize high-dimensional data in a two-dimensional space. It's particularly useful when we have data with many features or dimensions, like the butterfly wing measurements. T-SNE works by first computing a similarity measure between all pairs of data points, and then it maps each data point to a two-dimensional point in such a way that similar points are close together and dissimilar points are far apart. My teammates will introduce more details later.

Now, you might be wondering how t-SNE is different from other dimensionality reduction techniques, like PCA, which we’ve learned on class. While both t-SNE and PCA are methods for reducing the number of dimensions in a dataset, they do so in different ways.

PCA is a linear method that seeks to find the directions in the data that explain the most variance. It then projects the data onto these directions to reduce the number of dimensions. This can be useful for finding the most important features of the data, but it can also lose information about the relationships between different points.

T-SNE, on the other hand, is a nonlinear method that preserves the local structure of the data. It's particularly good at preserving clusters and patterns in the data, which can be very helpful for visualization and exploratory data analysis.

These two pictures are their performance on a mushroom classification dataset to determine the mushrooms are safe to eat or deadly poison. Based on what it shows, we can say compared to the performance of PCA, t-SNE performs better. The t-SNE algorithm clusters the poisonous and edible mushrooms without any overlap. PCA is unable to classify the mushrooms perfectly.

In summary, unlike PCA, t-SNE can be better applied to datasets with both linear and nonlinear well-clustered data, and produce more meaningful clusters. Although t-SNE is excellent at visualizing well-separated clusters, in most cases it is unable to preserve the overall geometric shape of the data. Therefore, both PCA and t-SNE have their advantages and disadvantages.