**Principal Component Analysis (PCA)** is a linear dimensionality reduction technique that identifies the most important variables in a dataset by projecting the data onto a lower-dimensional space while retaining most of the original information. PCA seeks to find a new set of variables that are linear combinations of the original variables, and that account for the most variance in the data.

**Non-Negative Matrix Factorization (NMF)** is a linear dimensionality reduction technique that factorizes a non-negative matrix into two non-negative matrices, where the product of the matrices approximates the original matrix. NMF is particularly useful for finding patterns in non-negative data, such as images and audio, and can be used for feature extraction or data compression.

**t-Distributed Stochastic Neighbor Embedding (t-SNE)** is a nonlinear dimensionality reduction technique that seeks to preserve the local structure of the data in a lower-dimensional space. t-SNE models the high-dimensional data as a probability distribution and maps it to a lower-dimensional space using a t-distribution to better represent the distances between points. t-SNE is useful for visualizing high-dimensional data in a lower-dimensional space while maintaining the relationships between the points.

Overall, PCA, NMF, and t-SNE are powerful techniques for reducing the dimensionality of high-dimensional datasets, while retaining important information and structure. Each technique has its own strengths and weaknesses, and the choice of which to use depends on the specific problem and data at hand.

**Principal Component Analysis (PCA):**

PCA finds a set of orthogonal vectors, called principal components, that describe the directions of maximum variance in the data. The principal components are ordered by the amount of variance they explain, and can be used to project the data onto a lower-dimensional space.

1. Calculate the covariance matrix of the data X
2. Calculate the eigenvectors and eigenvalues of the covariance matrix
3. Order the eigenvectors by their corresponding eigenvalues
4. Choose the k eigenvectors corresponding to the largest eigenvalues to form the new basis
5. Project the data onto the new basis to obtain the transformed data X'

Non-Negative Matrix Factorization (NMF):

NMF factorizes a non-negative matrix X into two non-negative matrices W and H, where X ≈ WH. The factors W and H represent the basis and coefficients of the decomposition, respectively.

1. Initialize W and H with non-negative values
2. Update W and H iteratively using multiplicative update rules until convergence:
3. H\_new = H \* (W.T @ X) / (W.T @ W @ H + epsilon)
4. W\_new = W \* (X @ H.T) / (W @ H @ H.T + epsilon)
5. Normalize W and H to have unit norm

**t-Distributed Stochastic Neighbor Embedding (t-SNE):**

t-SNE is a nonlinear dimensionality reduction technique that aims to preserve the local structure of the data in a lower-dimensional space. t-SNE models the high-dimensional data as a probability distribution and maps it to a lower-dimensional space using a t-distribution to better represent the distances between points.

1. Calculate the pairwise distances between all points in the high-dimensional space
2. Compute the probability distribution q\_i|j that point i would pick point j as its neighbor, using a Gaussian kernel:
3. q\_i|j = exp(-||x\_i - x\_j||^2 / 2sigma\_i^2) / sum\_k exp(-||x\_i - x\_k||^2 / 2sigma\_i^2)
4. Compute the probability distribution p\_i|j that point i would pick point j as its neighbor, using a t-distribution with one degree of freedom:
5. p\_i|j = (1 + ||x\_i - x\_j||^2)^-1 / sum\_k (1 + ||x\_i - x\_k||^2)^-1
6. Optimize the low-dimensional embeddings by minimizing the Kullback-Leibler divergence between the two distributions:
7. KL(P || Q) = sum\_i sum\_j p\_i|j log(p\_i|j / q\_i|j)
8. Update the low-dimensional embeddings using gradient descent, using the gradient of the KL divergence with respect to the low-dimensional coordinates. Repeat until convergence.