1. Reasons and Disadvantages:

Reasons:

* Reduce computational cost: Processing and analyzing high-dimensional data can be expensive.
* Improve model performance: High dimensionality can lead to overfitting and reduced accuracy.
* Enhance data visualization: Complex data is difficult to visualize in high dimensions.
* Gain insights: Uncover hidden patterns and relationships easier.

Disadvantages:

* Information loss: Valuable information might be discarded during reduction.
* Choice of technique: Selecting the right method depends on the data and task.
* Interpretability: Reduced features might be harder to interpret.

2. Dimensionality Curse:

As dimensionality increases, the volume of data needed to maintain the same level of accuracy grows exponentially. This makes learning, analysis, and visualization increasingly challenging in high dimensions.

3. Reverse Dimensional Reduction:

Generally, no perfect way to completely reverse dimensionality reduction exists. Once information is discarded, it's often irretrievable. However, some techniques offer partial reconstruction:

* PCA: Project data back onto the original features using the retained principal components. Approximates original data but loses discarded information.
* Autoencoders: Neural networks trained to reconstruct input from low-dimensional representations. Can learn more complex relationships but reconstruction might not be perfect.

4. PCA for Non-linear Data:

PCA assumes linearity. While it can still reduce dimensions, it might not capture the true structure of highly non-linear data. Consider non-linear methods like:

* Kernel PCA: Projects data into a higher-dimensional space to find linear separability, then reduces dimensions in that space.
* Manifold learning: Assumes data lies on a lower-dimensional manifold embedded in the high-dimensional space and seeks to recover that manifold.

5. PCA Explained Variance Ratio:

A 95% explained variance ratio means PCA retains 95% of the data's variance in the new, lower-dimensional representation. The exact number of dimensions depends on the data distribution and eigenvalues. You can determine it using software libraries that implement PCA.

6. Choosing the Right PCA Variant:

* Vanilla PCA: Suitable for smaller datasets and initial exploration.
* Incremental PCA: Processes data in batches, useful for very large datasets.
* Randomized PCA: Faster approximation, especially for high dimensions.
* Kernel PCA: Applicable for non-linear datasets, often computationally expensive.

Consider factors like dataset size, computational constraints, and the need for linearity when choosing a method.

7. Assessing Success:

Metrics depend on the application:

* Reconstruction error: How well the reduced data reconstructs the original data (PCA).
* Clustering performance: How well the reduced data separates clusters (clustering tasks).
* Classification accuracy: How well the reduced data helps model performance (classification tasks).

8. Chaining Dimensionality Reduction Techniques:

Yes, it's possible. If one method doesn't capture all relevant information, chaining different techniques can be beneficial. For example, you could use:

* PCA followed by kernel PCA: Capture linear and non-linear structures sequentially.
* Autoencoders in conjunction with PCA: Use autoencoders for complex non-linearities and PCA for further compression.