1. What are the key reasons for reducing the dimensionality of a dataset? What are the major disadvantages?

Ans: Disadvantages of Dimensionality Reduction:

- It may lead to some amount of data loss.
- PCA tends to find linear correlations between variables, which is sometimes undesirable.
- PCA fails in cases where mean and covariance are not enough to define datasets.

2. What is the dimensionality curse?

Ans: The curse of dimensionality basically means that the error increases with the increase in the number of features. It refers to the fact that algorithms are harder to design in high dimensions and often have a running time exponential in the dimensions.

3. Tell if its possible to reverse the process of reducing the dimensionality of a dataset? If so, how can you go about doing it? If not, what is the reason?

Ans: No, dimensionality reduction is not reversible in general.

4. Can PCA be utilized to reduce the dimensionality of a nonlinear dataset with a lot of variables?

Ans: Well it Depends on dataset. If it is comprised of points that are perfectly aligned, PCA can reduce the dataset down to 1 dimension and preserve 95% of the variance.

5. Assume you're running PCA on a 1,000-dimensional dataset with a 95 percent explained variance ratio. What is the number of dimensions that the resulting dataset would have?

Ans: If I perform PCA on a 1,000-dimensional dataset, setting the explained variance ratio to 95%. In this case roughly 950 dimensions are required to preserve 95% of the variance. So the answer is, it depends on the dataset, and it could be any number between 1 and 950.

6. Will you use vanilla PCA, incremental PCA, randomized PCA, or kernel PCA in which situations?

Ans: The following are the scenarios where the following are used:

- Vanilla PCA: the dataset fit in memory
- Incremental PCA: larget dataset that don't fit in memory, online taks
- Randomized PCA: considerably reduce dimensionality and the dataset fit the memory.
- kernel PCA: used for nonlinear PCA

7. How do you assess a dimensionality reduction algorithm's success on your dataset?

Ans: By doing PCA, it is a good choice for dimensionality reduction and visualization for datasets with a large number of variables.

8. Is it logical to use two different dimensionality reduction algorithms in a chain?

Ans: Indeed, it often make any sense to chain two different dimensionality reduction algorithms. A common example is using PCA to quickly get rid of a large number of useless dimensions, then applying another much slower dimensionality reduction algorithm, such as LLE.