1 Principal Component Analysis

Recommended applications of PCA

- **Data visualization:** Reduce data to 2D (or 3D) so that it can be plotted.

  This is a good use of PCA, as it can give you intuition about your data that would otherwise be impossible to see.

- **Data compression:** Reduce the dimension of your data, so that it takes up less memory / disk space.

  If memory or disk space is limited, PCA allows you to save space in exchange for losing a little of the data's information. This can be a reasonable tradeoff.

Inappropriate uses of PCA

- **Data visualization:** To take 2D data, and find a different way of plotting it in 2D (using k=2).

  You should use PCA to visualize data with dimension higher than 3, not data that you can already visualize.

- To get more features to feed into a learning algorithm.

  PCA will reduce the number of features, not expand it.